THE USE OF REGRESSION MODELS IN ANALYTICAL REVIEW JUDGMENT: A LABORATORY EXPERIMENT

BY

H. FRANCIS BUSH

A DISSERTATION PRESENTED TO THE
GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1989

ACKNOWLEDGEMENTS

I acknowledge the significant contributions of my dissertation committee members: William F. Messier, E. Michael Bamber, and Pejaver V. Rao. I especially thank William F. Messier who provided excellent guidance and demonstrated great patience as the chairman of the committee. I appreciate the accounting faculty for their various contributions in shaping the experiment, in designing the methodology and in offering constructive criticism and support.

I am grateful for the generous financial support of both the Fisher School of Accounting and the Deloitte, Haskins and Sells Education Foundation. I appreciate the auditors for participated in the experiment for their time and effort. I acknowledge the considerable contribution of John Carey and David Hunerberg, who made the arrangements for their firms to distribute and collect the experimental data.

An individual's success is the product of the support and assistance of his family and friends. I thank my wife, Felicia, and my children, Adina and Alex, for their understanding and support.

Lastly, I would like to dedicate this dissertation in the memory of my grandmother, Ottilie Kothe Lane.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

THE USE OF REGRESSION MODELS IN ANALYTICAL REVIEW JUDGMENT: A LABORATORY EXPERIMENT

BY

H. Francis Bush

May 1989

Chairman: William F. Messier
Major Department: Fisher School of Accounting

A laboratory experiment was conducted to investigate the relative merit of regression models as aids to analytical review judgment. The effects of environmental factors were also explored. The purpose of this study was to assess the effects of two sets of factors on the auditors' judgment. The first set of factors related to the decision approach and to the auditors' firms technology. The second set of factors related to the size of the variance of random noise and the size of the seeded error. Ninety-one auditors from two Big Eight CPA firms were asked to complete two tasks. Auditors from each firm were assigned to a different decision approach. The first task required each auditor to estimate the current monthly sales

figures and the bounds of the non-investigation region for four companies. Each company was created using simulated data from macroeconomic sources. The simulation used two levels for the variance of random noise and two levels for the size of the seeded error. The second task required the auditors to review the reported monthly sales figures, to indicate if the reported values were within the bounds of the non-investigation region, and to indicate their confidence in the implied decision about additional investigations. The results indicated that (1) the use of regression models as an aid to analytical review judgment for auditors from the unstructured firm was associated with an increase in their forecasting accuracy, a decrease in their type I error rate and an increase in their type II error rate, (2) the use of regression models as an aid to analytical review judgment for auditors from the structured firm was associated with a decrease in their type I error rate, while their forecasting accuracy and type II error rates remained unchanged, and (3) the use of regression models was associated with an increase in the confidence of auditors from the unstructured firm in their implied decision not to investigate further.

CHAPTER I

INTRODUCTION

The auditing profession is constantly seeking improved and more efficient auditing procedures. The use of analytical procedures during the audit can play a critical role in identifying areas that can have a significant impact on the audit. These procedures can also assist in effective and efficient audit planning (Felix and Kinney, 1982). This dissertation will investigate one possible improvement in the application of analytical procedures during the planning stages of the audit.

Analytical procedures are audit tests that provide evidence of the validity and reasonableness of accounting transactions and balances based on "plausible relationships among both financial and nonfinancial data" (A.I.C.P.A., p.1). The Auditing Standards Board indicates that analytical procedures may be performed for any of three purposes:

a. To assist the auditor in planning the nature, timing, and extent of other auditing procedures

b. As a substantive test to obtain evidential matter about particular assertions related to account balances or classes of transactions

c. As an overall review of the financial information in the final review stage of the audit (p. 2).

In applying analytical procedures three sequential judgments are required: extrapolation, comparison, and choice. The first judgment requires the auditor to establish a range of reasonable values for the account balance. This range is established by reviewing past trends in the account balance and finding relationships with other accounts. The second judgment requires the auditor to assess the significance of any difference between the reported account balance and the predicted range. Finally, based on the comparison, the auditor chooses the extent and timing of additional audit work.

The auditing literature indicates that there are two basic approaches to the first two judgments: judgmental and mechanical. The judgmental approach requires the auditor to subjectively extrapolate the range of reasonable values and to compare that range to the client's reported value. The mechanical approaches vary from simple rules of thumb to formal statistical models. Simple rules of thumb base the decision about additional audit work on a small set of cues, such as the presence of a material error in the previous audit. The formal statistical models derive a range by using a formula and subsequently applying a decision rule about additional audit work. Although both approaches are prevalent in practice, there is potential for a third and more effective approach. Recent research in other

disciplines (e.g., forecasting and medical diagnosis) indicates a third alternative that combines judgmental and mechanical approaches into a single decision approach which can outperform either approach individually (Sawyer, 1966). The purpose of this dissertation is to investigate the relative merits of combining statistical models for analytical procedures with auditors' judgment during the planning stage of the audit.

Motivation

This research was motivated by the following considerations. First, economic pressures requiring auditors to increase their efficiency and to hold the line on costs have drawn increased attention to the potential of using analytical procedures during the planning stage of an engagement (Blocher et al., 1983). Second, studies on auditors' analytical review judgments indicate that auditor judgment is biased (e.g., Biggs and Wild, 1985). The bias in auditors' judgments may in turn adversely affect the quality of their decisions (Libby, 1981). Third, research on analytical procedures has focused mainly on the effectiveness of mechanical approaches (e.g., Kinney and Salamon, 1983). Fourth, research in related disciplines, such as forecasting and medical diagnosis, indicates that combined approaches potentially may outperform either judgmental or mechanical approaches alone. Fifth, the complexity of the audit task can affect auditors' ability to

extrapolate current unaudited account balances (Biggs and Wild, 1985). Sixth, the audit firm's technology and training may affect the application of any audit procedure (Cushing and Loebbecke, 1986).

Economic Factors

The first consideration is the economic pressure facing the auditor in the current, more competitive environment. Arrington et al., (1983) have noted that the demands on the auditor are increasing. Clients are becoming larger; business systems are more complex; there is more regulation; visual audit trails are decreasing; and operating costs are increasing. Competition requires auditing firms to find ways to increase the efficiency and effectiveness of the audit. Analytical procedures during the planning stage of the audit are an important way to meet this challenge. During the planning phase, analytical procedures provide assistance by directing attention to account balances and transaction cycles with potential problems or errors (Kinney, 1983). Identification of such items assists the auditor in identifying critical areas, in allocating audit effort, and in choosing an initial set of audit procedures.

Bias in Analytical Review Judgment

The second consideration is the presence of bias in auditors' judgment. The presence of bias in auditors' judgments requires further study to determine the effects on the quality of auditors' decisions and to investigate

procedures to minimize these effects (Libby, 1981). Kinney and Uecker (1982) demonstrated that auditors' analytical review judgments were influenced by the presence of clients' reported values. Further, Biggs and Wild (1985) indicated that auditors' forecasts tended to underestimate non-linear trends. Combining mechanical approaches with analytical review judgments may diminish these biases.

Research on Analytical Procedures

The third consideration pertains to the research on analytical procedures. Kinney (1978) has compared various statistical models, such as regression and ARIMA, to determine which has the greatest accuracy in signalling significant changes in reported account balances. However, research has not evaluated the relative efficiency of judgmental approaches versus statistical models for signalling significant changes in reported account balances. Nor have the possible merits of combining statistical models with auditor judgment been examined.

Research in Related Disciplines

Research in medical diagnosis and forecasting has investigated individual methods of decision making (judgmental and mechanical). This research has analyzed the merits of each approach as well as comparative merits of both (Sawyer, 1966). Libby (1981) discusses the success of adapting the approaches of other fields to auditing (e.g., Fryback and Thornbury, 1978). The success of these

adaptations, and the parallel nature of the judgmental tasks in medical diagnosis and forecasting, and the judgmental task in analytical procedures, suggest extending this research to analytical review judgments.

Complexity of the Judgmental Environment

Increasing or decreasing trends, linear or non-linear trends, the size of the error and the seeding of the error, are among the many possible environmental components affecting auditors' judgments. Biggs and Wild (1985) reported that the direction of the trend, and the nature of the underlying function, affected the auditors' abilities to extrapolate current sales figures. Knechel (1984) has demonstrated, in a simulation study, that the size of the seeded error and the pattern of seeding the error affected the relative effectiveness of various mechanical approaches.

Audit Firm Technologies

Sawyer (1966) indicated that clinicians' training and affiliation may bias their approach to decision making. There is evidence that different Big Eight CPA firms have different thresholds when making decisions (e.g., Gaumnitz et al., 1982 and Joyce, 1976). Specifically, Cushing and Loebbecke (1986) discuss two polar approaches to auditing, structured and unstructured, which are used by Big Eight CPA firms. The audit firm's technology must be considered when evaluating the merits of a combined approach.

Summary

This dissertation will investigate the relative merits of the judgmental approach used alone and the judgmental approach aided by a statistical model. This topic was motivated by six considerations: economic factors, research on auditors' analytical review judgments, research on analytical procedures, research in related disciplines, the complexity of the judgmental task and the firm's audit technology.

This combined approach is expected to increase the effectiveness and efficiency of analytical procedures. The combined approach should reduce bias in the audit judgment. The use of this approach should be more successful with structured firms. The relative effectiveness of the combined approach should vary based on the level of complexity of the audit task. The results of this dissertation should direct the auditors' choice of analytical procedures based on their desire to control type I or type II errors. Finally, this experiment extends the study of the use of analytical procedures as suggested by Felix and Kinney (1982) for directing audit activities.

With this objective in mind, the remainder of the dissertation is organized in the following manner. Chapter II will review the literature from which the specific hypotheses were derived. The methodology for this research

is discussed in Chapter III. The results will be presented in Chapter IV. In closing, Chapter V will discuss conclusions and limitations.

CHAPTER II

LITERATURE REVIEW

Introduction

A literature review is required to clarify motivations, to reveal additional research issues and to focus the hypotheses statements. When dealing with the issues surrounding analytical procedures and audit judgments, six areas of research were explored: analytical procedures in auditing, human information processing research in auditing, statistical models as analytical procedures, forecasting, judgment and decision aids and audit technology. Although these areas are discussed separately, they are interrelated. Their relationships and how they impact the hypotheses will be discussed later in the chapter.

First, the potential for analytical procedures to increase the efficiency and effectiveness of the audit is demonstrated both in professional pronouncements and in studies on their applications (e.g., Holder, 1983, and Blocher et al., 1983). Second, research on human information processing indicates that heuristics affect auditors' analytical review judgments which may hinder the usefulness of judgmental analytical procedures. Third, research investigating the use of statistical models

recommends combining statistical models with auditors' judgment, illustrates their potential as analytical procedures and highlights the importance of environmental factors on the relative performance of each procedure. Fourth, the forecasting literature suggests the combination of statistical models with judgment to improve forecasting judgment. Fifth, research on decision aids suggests that combining statistical models with judgment, outperforms either mechanical models or judgment alone. Sixth, recent findings about audit firm technologies suggest that current research on audit procedures need to include this factor.

Analytical Procedures in Auditing

The significance of the research questions is based on the importance of analytical procedures to the audit. The purpose of this section is to describe analytical review judgment, to demonstrate its importance to the audit and to describe its application in practice.

Auditors' analytical review judgments are centered on forming expectations of the audit period's account balances, ratios, percentages, et cetera. Subsequently, the auditors' attention is directed to account balances with significant differences between expectations and client's reported values. Lastly, the auditor selects the timing, extent, and nature of other audit procedures. The results of analytical procedures can play a critical role in identifying areas with significant impact on the audit.

Analytical procedures have been identified as an important element in the audit process (A.I.C.P.A., 1988). Statement on Auditing Standards Number 56 noted that "Understanding financial relationships is essential in planning and evaluating the results of analytical procedures, and generally requires knowledge of the client and the industry or industries in which the client operates" (p. 1).

Holder (1983) assessed the frequencies of various techniques and the economic phenomena to which analytical procedures are applied. He suggested that (1) analytical procedures were most beneficial when a limited number of financial statement accounts were included, (2) trend analysis was used by approximately 60 to 70 percent of the subjects, and (3) areas of interest included account balances, liquidity, profitability/capital maintenance, and operational and financial structure.

Blocher, Esposito and Willingham (1983) reported the following findings about auditors' choices between analytical procedures and test of details. First, the use of analytical procedures increased as the audit budget decreased. Second, auditors tended to anchor on the previous year's audit program's scope of test of details. Third, the auditors preferred trend analysis and operating data for analytical procedures.

Analytical procedures' pivotal role in audit planning, and the success of analytical procedures to indicate errors, (see Hylas and Ashton, 1982, and Kinney, 1979) demonstrate the potential to meet the auditors' economic demand to increase the effectiveness and efficiency of the audit.

Human Information Processing Research in Auditing

The use of heuristics in auditors' judgments has been noted in the research on human information processing in auditing. The use of heuristics can cause systematic errors. These heuristics include anchoring and adjustment, overemphasis of linear trends and conservatism. This section will demonstrate that these heuristics affect auditors' analytical review judgments.

Anchoring and Adjustment

Using the anchoring and adjustment heuristic, the decision maker selects an initial value, the anchor, and then adjusts away from the anchor when additional information becomes available. The anchor may be selected based on past experience, initial computations, or some outside source (Tversky and Kahneman, 1974). Typically, the adjustment is made in the appropriate direction, but is not sufficient to match the judgment of a Bayesian model.

Kinney and Uecker (1982) investigated the effects of anchoring and adjustment on auditors' analytical review judgment. They asked auditors to provide estimates of the bounds of the non-investigation region of the gross profit

percentage. Each auditor received two years of previously audited data and the current unaudited information. They manipulated the current unaudited balances and ratios to provide two anchors, one above and one below the two previous years' audited gross profit percentages. Auditors from five Big Eight CPA firms were randomly assigned to each group.

Kinney and Uecker demonstrated the following two effects of the anchoring and adjustment heuristic on auditors' analytical review judgment. First, the mean responses for both boundaries were significantly lower for the group receiving the low anchor. Second, the percentages of responses which excluded the anchor from the non-investigation region differed significantly. Sixty-one percent of the group receiving the low anchor excluded the anchor while 81 percent of the group receiving the high anchor excluded the anchor.

Although, Kinney and Uecker demonstrated the effects of the anchoring and adjustment heuristic on auditors' analytical review judgments, several limitations should be considered. First, Kinney and Uecker did not provide a control group, a group which did not receive unaudited values. Second, the low anchor was consistent with a downward trend in the gross profit percentages, while the high anchor was consistent with an alternating pattern. Third, the cases were fictitious and may not have been

representative of any actual company. Fourth, Kinney and Uecker did not provide a control for differing auditors' thresholds for the term "materially misstated."

Biggs and Wild (1985) made the following extensions to Kinney and Uecker's work. First, they included control groups that did not receive any current unaudited information (anchor) and which were used to measure the effects of the presence of the current unaudited information. Second, Biggs and Wild gave some groups five years of previous audited data and other groups two years of data. Third, "a constant noninvestigation interval (95%) was established for all subjects (p. 613)." These extensions were used to further investigate the effects of anchoring and adjustment on auditors' analytical review judgment.

Biggs and Wild's findings can be discussed in two general categories: comparing the effects between groups which received two years of audited data and comparing the effects between groups which received two years with the groups which received five years of audited data.

Using two years of audited data, Biggs and Wild demonstrated the following points. First, the responses were consistent with the findings of Kinney and Uecker. Second, the absolute differences between the group receiving the high anchor and the control group were significant, while the absolute differences between the

group receiving the low anchor and the control group were not significant. Third, the control group excluded the high unaudited value significantly more frequently than those in the group receiving the high anchor, while the control group did not exclude the low unaudited value significantly more frequently than those in the group receiving the low anchor. Biggs and Wild concluded that auditors receiving two previous years of audited data were anchoring more on the high unaudited value or that the auditors were not distinguishing the unaudited nature of the high anchor.

Comparisons with the groups receiving five previous years of audited data indicated the following. First, the responses of auditors receiving the longer data series were significantly different from the respective groups receiving the shorter data series. Second, the widths of the confidence interval were not significantly narrower for the groups receiving longer data series. Third, although the bias of the anchors was smaller and the significance was less for auditors receiving the longer data series, the heuristic persisted in influencing the auditors' judgment. Fourth, the number of implied investigations (i.e., the anchor is outside of the bounds of the non-investigation region) was not significantly different for groups receiving the longer data series. Fifth, there was evidence that auditors identified the decreasing gross profit percentages as a trend.

Based on these findings Biggs and Wild reported the following conclusions. First, auditors apparently used the additional years of information in making their decisions. Second, confidence as measured by the size of the non-investigation region was not affected by the presence of the additional years of information. Third, the effects of anchoring and adjustment were mitigated but not eliminated.

Although Biggs and Wild demonstrated the persistence of the effects of the anchoring and adjustment heuristic, several limitations of their work need to be addressed. First, Biggs and Wild's experiment was based on fictional data. The anchors, which were gross profit percentages with the same level of sales, were created simply by varying only cost of sales. The higher percentage was derived by decreasing cost of sales and increasing sales which might have created a more suspicious event. Second, the anchors were extreme values. In each case the unaudited gross profit percentage was outside of the range of previously audited percentages. Third, the auditors responded in terms of percentages relating two significant income account balances. Auditors might have been more familiar with responses using dollar amounts for each account balance separately. Fourth, confidence is measured indirectly. The width of the non-investigation region is affected differently by the desire to detect errors and the desire to reduce audit effort. These issues and limitations must be

considered when interpreting the results which Biggs and Wild discovered in their experiment.

Anchoring and adjustment is a heuristic which affects the auditor's analytical review judgment. It can be moderated with additional information but may not be eliminated. The previous studies on anchoring and adjustment motivate three facets of this research study. First, it is necessary to acknowledge and minimize the effect of the anchoring and adjustment heuristic. Second, the relative merits of analytical procedures need to be investigated using cases that more closely simulate actual data. Third, auditors' confidence in their judgments should also be included in the study.

Overemphasis of Linear Trends

Analytical review judgments require auditors to extrapolate time series data. Decision makers overemphasize linear trends either by simplifying non-linear patterns or perceiving linear patterns where no pattern exists.

Eggleton (1982) demonstrated the difficulties students had extrapolating trends and identifying random and alternating patterns. Each student was presented with 12 numbers, representing monthly cost figures, and was asked to provide the next predicted value and the bounds of a credible interval. Eggleton concluded that individuals (1) distinguished trends, (2) perceived alternating patterns as random, and (3) imposed "lawfulness" on random patterns,

which increased with the size of the variance and the amount of time allowed.

The overemphasis of linear trends affects the auditors' ability to extrapolate the current period's account balance, ratio, et cetera. Similar to anchoring and adjustment, overemphasis of linear trends may reduce the effectiveness of analytical procedures during the planning stages of the audit. This heuristic, however, is often applied with a second heuristic, conservatism.

Conservatism

Conservatism occurs when the decision maker avoids extreme values. Applying the conservative heuristic causes the auditor to underestimate the trend in the data. Wickens (1984) offered two causes for the conservatism in extrapolating serial data. First, individuals may have a tendency to avoid extreme values. Second, individuals may have learned from the actual environment that there are limits to any expansion or trend. Conservatism along with overemphasis of linear trends may create difficulty for auditors in making the first judgment required in using analytical procedures.

In a second experiment, Biggs and Wild (1985) investigated auditors' ability to extrapolate various trends in annual sales data. Each auditor was provided with seven years of audited revenue data and told that there was no significant change in the underlying relationships. Each

auditor received a case representing one of six treatments. The treatments were generated by crossing deterministic linear, exponential, and logarithmic functions with increasing and decreasing trends.

Biggs and Wild results were as follows. Consistent with the effects of conservatism, all the increasing functions were underestimated. Further, the auditors underestimated the exponential function significantly more than the linear and logarithmic functions. The results were similar with the decreasing trends except for the linear trend. The auditors' extrapolated values for the decreasing linear function were not significantly biased.

Conservatism, combined with overemphasis of linear trends, has been demonstrated to hinder auditors' ability to extrapolate trends in account balance. Wickens (1984) summarized three causes for the difficulty in extrapolating time series. First, linear representations were less complicated and required less cognitive strain. Second, individuals expressed a tendency to avoid extreme values. Third, individuals may have learned from the actual environment that limits are imposed on expansion; individuals may transfer this notion to their estimation procedures. These heuristics pose a threat to the effectiveness of analytical procedures because the first judgment in any analytical procedure requires the auditor to

extrapolate account balances, ratios, et cetera from trends in the data.

Summary

Heuristics affect auditors' analytical review judgment. First, anchoring on current unaudited reported values affects the auditors' judgment. Second, overemphasis of linear trends occur when individuals reduce complex patterns or attempt to impose patterns on random order. Third, conservatism causes auditors to underestimate their extrapolated values which would be used for comparison with the client's reported value. If analytical procedures are to have an effective impact on audit planning, auditors must understand and overcome the biases created by the use of these heuristics. One such approach which has been the topic of research, is the use of statistical and other mechanical approaches as analytical procedures.

Statistical Models as Analytical Procedures

The prominence of regression models in the research on mechanical analytical procedures has been demonstrated in detail by Arrington et al. (1983). This section reviews the role of regression models in the research on mechanical approaches as analytical procedures. The purpose of this review is to (1) demonstrate the relative merit of regression models as analytical procedures and (2) develop methodological issues used in the experiment of this dissertation. To examine the relative merits of regression

models, it is necessary to survey studies using both simulated data and actual data.

Regression Models in Simulation Studies

Kinney and Salamon (1982) and Knechel (1984) reported on two studies which used simulated data. Kinney and Salamon investigated the relative merits of different investigation rules used in conjunction with regression models. The environmental parameters which they varied were the size of material error and the size of the seeded error. Relative performance was measured in terms of type I errors (the procedure signals the need for additional audit work when there is no material misstatement in the account) and type II errors (the procedure fails to signal the need for additional audit work when there is material misstatement of the account). Kinney and Salamon concluded that amending the Deloitte, Haskins and Sells' STAR approach with an ex post annual balance check would provide both an efficient and effective analytical procedure.

Knechel (1984) expanded Kinney and Salamon's work by including additional models and investigation rules, including additional underlying populations, and considering cost factors. Knechel included submartingale models along with the regression models. The underlying populations were simulated using the environmental parameters, such as the size of the seeded error in relation to materiality and the distribution of the error across the months. To allow

auditors to combine the findings of the study with loss functions, Knechel presented the relative performances by graphing type I errors versus type II errors for each simulated population.

Knechel demonstrated two results. First, in each population the regression models were consistently on the frontier of the graph of the error rates. Second, the rate of type I and type II errors were affected by the size of the seeded error in relation to materiality and the distribution of the error across the months of the audit period.

Regression Models and Actual Data Studies

Actual data studies obtain data from public records as well as from individual companies. Both Kinney (1978) and Lev (1980) based their studies on data from public record, while Akresh and Wallace (1981) based their study on an individual utility company.

Kinney (1978) investigated the relative predictability of time series models, regression models, and submartingale models as analytical procedures. Revenue data for 15 years for six railroad companies were obtained for the study. An independent variable was developed for each company by averaging the remaining five companies' revenue figures. Each analytical procedure's ability was measured in terms of the mean prediction error and the mean absolute error. Investigation rules, however, were not developed.

Kinney reported the following findings using mean absolute error. First, statistical models were superior to the submartingale models. Second, the transfer ARIMA functions, which used the most information available, outperformed the other models. Third, the regression models based on the independent variable were superior to the ARIMA functions based on autocorrelations. Fourth, the models using a 36-month base period generally outperformed the models based on 120 months.

Lev (1980) investigated the relative predictability of various regression models based on economic indices. Submartingale models were used as benchmarks and performance was measured by the mean absolute error. Lev reported the following findings. First, regression models based on the first difference of the national indices outperformed the models based on the national indices. Second, in general, the inclusion of industrial indices in the regression models improved the predictive ability of the models. Third, Lev reported that the 90 percent confidence intervals for the predicted values were generally narrow, representing low uncertainty. Fourth, several of the reported values were outside these intervals. Lev demonstrated the usefulness of regression models based on national indices across a variety of firms.

Akresh and Wallace (1981) investigated the usefulness and the application of regression models in a case study of

a utility company. Using a single case, Akresh and Wallace were able to develop multivariate regression models in an actual environment. These models were developed using variables suggested by auditors and variables suggested by step-wise regression.

In their study, Akresh and Wallace reported the following results. First, the step-wise regression models and the models suggested by the auditors, included many of the same variables. However, Akresh and Wallace stressed the importance of the auditors' judgment about the selection of variables for the model over the iterative step-wise regression procedures for selecting variables. Second, both models were successful at predicting revenues and other account balances.

Methodological Issues

These studies provided guidance in the development of the experimental task. The design factors which were suggested from this literature were (1) the development of case materials using simulation based on actual macroeconomic data, (2) the account for the judgmental task, (3) the performance measures, (4) the importance of manipulating environmental factors in the simulation, and (5) the time frame.

First, the instrument for the experiment was developed using simulations based on actual sales data and industrial indices because the sample cases could readily be developed

to incorporate various environmental factors and the "true value" would be available for performance measures. The generalizability of the results, however, are limited to the extent that the simulation captures the environment. Second, the sales account was included in each of these studies and was related to studies on auditors' analytical review judgment. Third, performance was measured using both the mean absolute percent error and error rates. Fourth, the simulation seeded the error over four out of twelve months, and the size of the seeded error was manipulated between one and two times materiality. Fifth, the time frame was five years including the prediction period.

Summary

Research on the relative merits of regression models as analytical procedures has made three contributions to this dissertation. First, Lev in his concluding comments suggested that regression models could assist auditors' analytical judgment. Second, the relative merits of regression models were demonstrated. And third, various methodological issues were discussed.

Forecasting

The forecasting literature was reviewed because the first judgment in analytical review requires the auditor to extrapolate the current unaudited values. As indicated above, the overemphasis of linear trends and the conservatism heuristics influenced the extrapolation

judgment of the current unaudited account balances. An improvement in the first judgment hopefully will improve the overall analytical review judgment.

Forecasting Method Studies

Studies in this area have focused on the abilities of numerous sophisticated forecasting methods using actual time series data.

Makridakis and Hibon (1979) investigated the relative abilities of 22 forecasting methods over 111 different time series. They used the mean absolute percent error and Theil's U statistic to rank each model's relative predictive abilities. They presented the following results. First, exponential smoothing had the overall best performance. Second, simple methods approximated sophisticated methods. Third, as random noise decreased, ARIMA's relative performance increased. Fourth, the amount of randomness in the series impacted the rankings of the forecasting methods. Fifth, Makridakis and Hibon recommended further investigation of combining statistical models with judgment.

To add more depth to this study, Makridakis, Anderson, Carbone, Fildes, Hibon, Lewandowski, Newton, Parzen and Winkler (1982) made the following extensions. They (1) included 1001 time series, (2) increased the number of comparisons, (3) increased the time horizon, (4) used back-forecasting to obtain the initial values, (5) used more measures of predictive accuracy, and (6) used experts in

each field to apply the models. Makridakis et al. (1982) reported the following observations. First, ARIMA models were the most time consuming, while regression models were less time consuming because they were completely computerized. Second, the effects of sampling from the time series affected the measures of accuracy. Third, the relative performance of each model was affected by the type of measure, the type of time series, and the length of the time horizon. Fourth, simpler models performed better with micro data or when randomness was higher. (They suggested that the sophisticated models "overestimated" their parameters.) Fifth, the average of six methods outperformed any of the individual six methods, which is consistent with the results reported by Makridakis and Winkler (1983).

The research on forecasting methods indicated several important factors for this experiment. First, the more sophisticated method may not have been appropriate with the micro-economic data (e.g., firm-level data). Second, the regression models were less costly than ARIMA models. Third, the amount of randomness, a factor to be manipulated in the simulation used in this dissertation, affected the relative performance of the forecasting methods. Fourth, the relative performance of the forecasting methods was also affected by the complexity of the time series. Fifth, the mean absolute percent error was also used as a measure of predictability. And sixth, several authors (e.g.,

Makridakis and Hibon, 1979) recommended the combination of statistical forecasting methods with judgment to increase accuracy.

Summary

The forecasting literature is related to the auditor's first decision in the analytical review judgment, predicting the current unaudited values. Review of this literature addresses three issues. First, improvements in this judgment is anticipated to improve auditors' analytical review judgment. Second, the experiment in this dissertation used various aspects of these studies. Third, the forecasting literature suggested the combination of statistical models and judgment. This recommendation leads to the review of the use of decision aids.

Judgment and Decision Aids

The combined approach of mechanical models and judgment has been successful in various areas (see Sawyer, 1963). To provide sufficient evidence that a combined approach will be useful to analytical review judgment, a review of this literature will provide background on the conceptual issues of human judgment, a framework of decision approaches, and findings on various approaches to human judgment in medicine and forecasting.

Conceptual Issues on Human Judgment

Before describing the findings of various approaches to judgment, it is necessary to provide a conceptual framework

of decision making. Libby (1981) discussed auditors' judgment in relation to behavioral decision theory. He provided the following general framework for decision making.

> In most decision-making situations, judgment about the environment must be made in the absence of direct contact with the objects or event to be judged. In such circumstances, 'most likely' judgments are formed on the basis of information or cues whose relationship to the object or event of interest are imperfect or probabilistic. That is, judgments and decisions are made under uncertainty about relationships between cues and events (p. 4).

Libby defined judgment as the process of predicting outcomes and their consequences, and decision making, as the process of evaluating the outcomes and their consequences to make the choice from the alternatives.

There are several considerations which suggest that mechanical approaches and judgments aided by mechanical approaches can outperform clinical judgment. Dawes (1971) noted that individuals, by their human nature, are affected by internal and external variables such as varying attention spans and environmental surroundings. Einhorn (1972) suggested that omission of important variables, non-optimal weights and cognitive strain, are factors which reduce experts' "global judgments." Edwards (1968), and Kahneman and Tversky (1973) discussed experts' inability to manipulate probabilities appropriately. Einhorn and Hogarth (1978) suggested that individuals overemphasize their positive hit rate because the individual may have a stake in

the outcome of an accepted decision and a rejected decision is usually not implemented.

The awareness of judgmental errors and the understanding of some of the causes, suggests the use of decision aids, mechanical tools or rules, which may prevent errors or reduce the degree of error. Several authors have discussed the basic requirements (see Pitz and Sachs, 1984, Wright and Ng, 1982 and de Dombal and Horrocks, 1974) and the advantages of decision aids (see Libby, 1981, Elstein et al., 1978 and Kendall, 1975).

De Dombal and Horrocks (1974) introduced the following requirements for a decision aid. De Dombal and Horrocks stressed that the expert is still the pivotal component in the process. The system must be both compatible for the expert and flexible for the variety of circumstances surrounding the task. To more accurately estimate statistical probabilities large data bases should be used. Wright and Ng (1982) stated that the goal of the decision aid is to maximize mechanical consistency and expert knowledge.

Kendall (1975) suggested that decision aids could reduce observer bias through standardization and increased reliability. Decision aids would also increase consistency in judgment, allowing experts to manage more information and apply probabilities more accurately. Decision aids could

also improve the organization of the information for decision making.

Before reviewing research studies comparing various judgmental approaches, it is necessary to consider another aspect of judgment; confidence in an individual's decision. In their review, Einhorn and Hogarth (1978) suggested four causes for increasing one's confidence. First, the cost of making an error might increase the efforts of the individual. Second, the environment does not always completely reveal itself. Third, individuals, including sophisticated users, have a tendency to fail to learn from disconfirming evidence. Fourth, an increase in the quantity of information has increased individuals' confidence, but not necessarily the accuracy of their judgments (see Kahneman and Tversky, 1973; Goldberg, 1968; and Oskamp, 1965).

Hilton (1980) addressed the effects of information on decision making from both a descriptive and a normative approach. The former explanation was based on information overload. Information overload describes the phenomenon in which subjects were observed to improve their decisions as information was provided, but only up to a point. Beyond some optimal point the decision making ability began to deteriorate.

From a rational choice point of view, Hilton explained this behavior as optimal. Individuals are aware of an

optimal number of cues that they can manage. As the number of cues increase, the individuals incorporate them up to a point at which the cognitive strain is more "costly" than the anticipated benefit. When there is an overabundance of cues, the individual must select a subset of cues, which in turn reduces the cognitive strain and the quality of decision making.

Research on the conceptual issues surrounding human judgment indicates the limitations and causes for the use of heuristics, the advantages of decision aids, the need for careful application of decision aids, and the need for an awareness of the effect of using a decision aid on the individual's confidence.

General Framework for Judgmental Approaches

Sawyer (1966) provided a framework to analyze decision approaches after reviewing 45 previous studies. His investigation revealed two distinct issues in the judgment process: cue collection and cue combination. Sawyer identified four approaches for cue collection (mechanical, judgmental, combined mechanical and judgmental, and either mechanical or judgmental) and two approaches for cue combination (judgmental and mechanical) (see Exhibit 1). In total, Sawyer defined eight decision approaches.

The first four of the approaches are obtained from crossing clinical and mechanical approaches with the issues of cue collection and cue combination. Pure clinical uses

CLASSIFICATION OF PREDICTION METHODS

Mode of data collection \ Mode of data combination	Clinical	Mechanical
Clinical	1. Pure clinical	2. Trait ratings
Mechanical	3. Profile interpretation	4. Pure statistical
Both	5. Clinical composite	6. Mechanical composite
Either or both *	7. Clinical synthesis	8. Mechanical synthesis

* Plus, for the clinical synthesis, the prediction of Method 2, 4, or 6; or, for the mechanical synthesis, the prediction of Method 1, 3, or 5.

Source: Sawyer (1966)

Figure 1.

judgment for both cue collection and cue combination; and pure statistical approaches uses mechanical methods for both cue collection and cue combination. Profile interpretation (clinical judgment) uses clinical judgment with mechanically derived cues. For trait rating, the judge clinically assesses the cues but mechanically combines them for a decision.

The remaining four approaches use both cue types. Clinical composition combines both types of cues using judgment. Mechanical composition combines both types of cues mechanically. Clinical synthesis (aided clinical judgment) uses a mechanical approach to combine cues. Experts then apply their judgment to make the final decision. Mechanical synthesis uses a mechanical rule to adjust the experts' judgments.

Using these definitions, Sawyer investigated the relative merits of each of these judgmental approaches. Using the results of 45 previous studies, he measured the percent of times each approach outperformed another approach. Sawyer's results were summarized as follows. First, no one approach was clearly superior. Second, no one pair of cue collection and combination was particularly well matched. Third, aided judgment outperformed other approaches in a greater percent of the comparisons than clinical judgment. Fourth, each approach outperformed another approach in at least 25 percent of the comparisons.

Sawyer's work impacted this dissertation by (1) demonstrating that no one approach was clearly superior, but indicated that aided judgment would probably outperform clinical judgment and (2) indicating that clinicians' education and training could affect the relative merits of any decision approach.

Research on Human Judgment Approaches

Research on the relative merits of various approaches to judgment has been conducted in many disciplines. This section will report on the findings of this research in medical diagnosis, forecasting and accounting.

Kendall (1975) conceived medical diagnosis as a judgmental process in which the expert matched a set of symptoms to a disease. For this matching process, Kendall defined reliability as the defining symptoms of a disease and validity as the correlation between symptoms (and the lack of symptoms) and the disease. Some cues are quantifiable, such as body temperatures, while others are qualitative, such as the headaches. Both medical diagnosis and analytical review judgments require experts to examine a set of cues of indeterminate certainty to assess the state of nature. The success of the judgment is the combination of accurately matching symptoms to the states of nature and determining the appropriate actions.

There are several other similarities between auditors' judgment and medical diagnosis. First, the experts' ability

will be affected by the quality of information about the symptoms, which are reported by the client, personal qualifications, the time allotted for the judgment, and the classification scheme. Second, inappropriate action can be extremely harmful to the client resulting in litigation for the expert. Third, experts try to be both efficient and effective in their judgments under uncertainty.

The findings on judgmental approaches in medical diagnosis have not produced a singularly preferred approach. Studies will be discussed involving clinical judgment, aided judgment, and mechanical approaches to judgment. These studies demonstrate the continuing need for research on judgmental approaches.

Leaper, Horrocks, Staniland and de Dombal (1972) and Einhorn (1972) provided evidence that mechanical judgment can outperform clinical judgment. Leaper et al. investigated the relative merits of three approaches to estimate the probabilities between symptoms and diseases relating to abdominal pain: a computer assisted system based on 600 actual case histories, a computer system using the clinicians estimates of the probabilities, and clinical judgment. Leaper et al. reported ratings of diagnostic accuracy of 79.7 percent, 82.2 percent, and 91.1 percent for the senior clinicians' judgment, the computer-assisted system based on the clinicians' estimates, and the computer-assisted system based on actual data, respectively.

Leaper et al. indicated that the ranking varied for specific diseases and "the effectiveness of the computer using estimates seemed to be related to the incidence of the disease under study" (p. 353).

Einhorn (1972) investigated the relative merits of global judgments, the use of multiple judges, and the mechanical combination of cues which were derived using judgment. Einhorn had three pathologists provide estimates of survival time and ratings on nine histological characteristics of 193 cases of Hodgkin's disease. Einhorn reported the following results. First, a disjunctive model was the best model for predicting the survival time. Second, a conjunctive model was the best model for predicting the pathologists' global judgment. Third, mechanical combination outperformed global judgment.

Einhorn suggested that the mechanical combination of cues outperformed the global judgmental diagnosis because the experts may not have use optimal weights. The mechanical combination may have relieved cognitive strain and the experts may have omitted important variables. He stated that, "these results were consistent with the notion of the nonlinear, noncompensatory model serving as a possible cognitive simplification mechanism" (p. 96).

Fryback and Thornbury (1978) and Gill, Leaper, Guillou, Staniland, Horrocks and de Dombal (1973) demonstrated that aided judgment outperformed clinical judgment. Fryback and

Thornbury (1978) investigated the relative abilities of clinical judgment, judgment aided by a numerical application of decision theory, and judgment aided by informal decision theory concepts, in selecting from two tests for diagnosing space-occupying renal lesion. Relative abilities were measured in terms of cost of testing, medical risks, and patient's discomfort and inconvenience. If the wrong test was chosen, it is assumed the clinician would become aware and administer the other test. Fryback and Thornbury concluded that the informal use of decision theory concepts reduced costs in terms of testing, medical risks, and patient discomfort. Further improvements were obtained when the numerical methods were applied.

Gill, Leaper, Guillou, Staniland, Horrocks and de Dombal (1973) investigated the abilities of computer-assisted judgment relative to clinical judgment. The computer-assisted judgment used pre-definitions and formalized approaches to reduce observer variations. Eight clinicians evaluated 552 actual patient histories involving abdominal pain. Gill et al. reported that computer-assisted judgment led to accuracy ratings of 91.5 percent to 95.3 percent based on the actual outcome.

These studies in the area of medical diagnosis indicated that mechanical and combined mechanical and judgmental approaches were superior to judgment alone. Several studies, however, in forecasting and accounting

(e.g., Lawrence, Edmunson and O'Connor, 1985, and Armstrong, 1983) imply that expert forecasting judgment might outperform statistical models.

Lawrence, Edmunson and O'Connor (1985) compared the abilities of novices, experts, and statistical models to extrapolate time series. Each group evaluated 111 time series which varied between micro and macro data, seasonal and nonseasonal series, and monthly, quarterly and annual data. Lawrence et al. indicated that judgmental extrapolations were no less accurate than statistical forecasting; that the judgmental extrapolation's standard deviation of the absolute percent error was smaller than the statistical forecasts; and that quantitative analytical tools may support judgmental forecasting.

Collins and Hopwood (1980) compared financial analysts' forecasts with basic time series models and specifically fitted Box-Jenkins ARIMA models. They reported that analysts' forecasts outperformed the statistical models; the significance between the analysts' forecasts and the models' forecasts decreased as the year-end became closer; accuracy for both approaches increased between quarters; and the analysts had fewer outliers. Collins and Hopwood hypothesized that the analysts were able to capture more information which was not captured by the models, such as a turn around after a strike.

Armstrong investigated the relative abilities of

management's forecasts, financial analysts' forecasts, and forecasting models. Armstrong reviewed 13 studies which produced 17 comparisons of forecasting models and judgmental (both management and financial analysts) forecasts. He reported that judgment was significantly more accurate using the mean absolute percent error for 14 of the comparisons. Armstrong questioned the results of the remaining three comparisons.

Armstrong also compared management's forecasts to financial analysts'. Reviewing five studies which allowed direct comparison between managements' and financial analysts' forecasts, Armstrong found their mean absolute percent errors were significantly different (15.9 percent versus 17.7 percent). Armstrong offered the following reasons for management outperforming the financial analysts. First, management had inside information and had more current information. Second, management influenced reported earnings. Third, management had control over performance and had set the goals of performance.

These studies demonstrated the ability of judgmental approaches to outperform mechanical approaches. However, the following cautions should be noted about the study by Lawrence et al. (1985): (1) the judgmental approaches were actually aided with graphs or tables; (2) many pair-wise comparisons were not significant; and (3) the illustration, which was provided in the appendix, had only 20 observations

which is less than the recommended number of observations for Box-Jenkins modeling (Chatfield, 1980). The lack of sufficient observations would bias the results against the mechanical forecasting. The studies by Collins and Hopwood (1980) and Armstrong (1983) used published forecasts, which might have been aided.

Summary

The literature on judgment and decision aids provided two contributions for this experiment. The first contribution developed a framework to investigate the use of decision aids for the analytical review judgment. The second contribution extended the scope of this study.

Mechanical aids and statistical models have the potential to assist analytical judgment because they reduce cognitive strain, they permit the individual to evaluate more information, they increase consistency, and the decision meets the recommendation of Edwards, Lindman and Phillips (1965) of being significant and repetitive, and of having quantitative inputs. The results of research on different judgmental approaches have been mixed. However, there is sufficient support in the literature for Lev's (1980) recommendation that the predictions of statistical models could enhance the auditors' analytical review judgment.

The study of judgment and decision aids also indicates two additional considerations. First,

confidence in an individual's decision can be affected by the amount of information or its presentation. Confidence, which was measured indirectly in previous studies on analytical review judgment, was measured directly. Second, the expert making the judgment may be affected by training and experience. The training of an auditor might be directly affected by the audit technology of the firm. This will be discussed in the next section.

Audit Technology

An individual's training and affiliation may affect the relative merit of any decision approach (Sawyer, 1966). Cushing and Loebbecke (1986) noted a difference in auditing approaches among the Big Eight CPA firms.

Cushing and Loebbecke reviewed the generally acceptable auditing standards in order to develop a model of the audit process. They investigated published materials of the Big Eight CPA firms and four national firms to develop individual firms' models of their auditing process to analyze audit structure and for cross-firm comparisons.

The classification was based on 18 elements of the firms' audit methodologies. Each element was ranked on a five point scale. Conflict was resolved by cross validation among the researchers. The classifications were highly structured, semi-structured, partially structured, and unstructured.

Cushing and Loebbecke defined four audit technologies

based on the structure of the firm's audit methodology.

> We define a structured audit methodology as a systematic approach to auditing characterized by a prescribed logical sequence of procedures, decisions, and documentation steps, and by a comprehensive and integrated set of audit policies and tools designed to assist the auditor in conducting the audit. A structured audit methodology is systematic in the sense that it views the audit process as a series of logical steps, often explicitly represented in the form of an audit process flowchart. The audit policies and tools used in a structured methodology are comprehensive in the sense that they cover the entire audit process, from the initial acceptance of the client to the ultimate issuance of the audit opinion. The audit tools are integrated by means of top-down approach to the design of the entire set of audit tools, such that the results of the audit may be incorporated into other audit tools used in planning later stages of the audit (p. 32).

Cushing and Loebbecke distinguished each classification by the firms in each category. The highly structured firms had the most developed approaches and extensive materials. They had pre-printed forms for guidance and explicit criteria for evaluating materiality and risk. Semi-structured firms' methodologies were not as developed as the highly structured firms. A more dramatic distinction was between the semi-structured and partially structured firms. The materials of the partially structured firms were interpreted as not having an overall concept of the audit process. Finally, the firms were classified as unstructured because either their materials lack analytical characteristics or their materials were viewed as incomplete.

Cushing and Loebbecke suggested that the structured approach was a response to greater competition, a need for line quality control, adverse effects of audit litigation, increasing complexity of data processing technology, and increasing complexity of the economy. The disadvantages included inflexibility, auditors' tendency to become mechanistic, increased vulnerability in audit litigation, cost to develop the approach, and the impact on the firms' personnel. In conclusion, Cushing and Loebbecke called for additional research on the effects of audit structure on audit judgment in terms of consensus, consistency, and the amount of time required for auditing judgments.

Kinney (1986) offered more evidence supporting the distinction between audit technologies among Big Eight CPA firms. Kinney investigated the effects of a firm's audit technology on the firm's interest in auditing standards. A firm's interest was measured by the votes or stated preference on six controversial issues of the Auditing Standards Board. An issue was considered controversial if three or more votes were cast against the exposure draft. Firms were classified in accordance with Cushing and Loebbecke's analysis by four members of the AICPA's statistical sampling subcommittee. These classifications were consistent with Cushing and Loebbecke's findings.

Kinney concluded that audit structure had the strongest relationship to voting patterns. Two additional variables

for the Big Eight CPA firms, which had a weaker but still significant effect, were staffing ratio and audit concentration. Finally, Kinney recommended that studies of "auditors' behavior with respect to procedures or reporting, or studying the organization of the profession ought to consider the potential effects of firms' technological preferences" (p. 88).

Research on audit firm technology has indicated that there are differences in the amount of structure in the audit process among the Big Eight CPA firms. This structure has been demonstrated to cause differences in the Big Eight CPA firms' responses to auditing standards. To test the relative merits of aided analytical review judgment, firms were selected representing polar positions on the structure of audit technologies. This was done because the research on judgment and decision aids indicated the importance of an expert's training. The effects of firms' audit technology will be considered in this study.

Hypotheses

This section will present the formal set of hypotheses. The issues being tested are: the relative merit of providing the predictions and standard deviations of regression models as an aid to analytical review judgment, and the effects of providing this information on the auditors' confidence in their decisions for audit planning.

Performance

The professional standards and the research on changes which were initiated by auditors indicate that analytical procedures can play a critical role in the allocation of audit resources during the planning stage. Research on human information processing in audit judgments and research on decision making indicates that heuristics and the inability to assess all the data have hindered the effectiveness and efficiency of analytical review judgment. The literature on judgment and decision aids further suggests that a decision aid would be appropriate for analytical review judgments. First, the analytical review judgment is significant and repetitive, and includes quantitative data. Second, a statistical model could serve to organize the quantitative information consistently and could reduce auditors' cognitive strain. Third, aided extrapolation judgments have outperformed forecasting models.

The first set of hypotheses compares the performance of aided and clinical analytical review judgment. The aided judgment will receive the predictions and standard deviation from regression models. Regression models have been shown to be a major competitor in the studies using statistical models as analytical procedures. Aided judgment should be enhanced by combining the advantages of mechanical cue combination and clinical judgment.

H_{1a}: Auditors using the decision aids will outperform auditors using clinical judgment alone.

The literature on judgment and decision aids indicates that experts' training may affect their ability to use a judgmental approach. The recent literature on audit firm technology indicates that the Big Eight CPA firms vary in the extent of structure in their audit processes. The use of regression models will be more compatible with firms which apply a more structured approach. The use of regression models may be less compatible with firms which apply a less structured approach.

H_{11a}: The aided analytical review judgment will outperform analytical review judgment more for auditors from the more structured firm than for auditors from the less structured firm.

Simulation studies on the use of statistical models as analytical procedures, and studies on forecasting, indicate that environmental factors can affect the relative abilities of judgmental approaches. Specifically, this experiment will manipulate two environmental factors, the size of the variance of random noise and the size of the seeded material error.

H_{12a}: Environmental factors (the size of the variance of random noise and the size of seeded error) which increase the difficulty of the judgment will increase the relative performance of aided analytical review judgment over clinical analytical review judgment.

Confidence

Confidence was a concern which was brought out by two of the disciplines reviewed. First, the literature on human

information processing research in auditing noted that confidence was an important variable in the auditor's analytical review judgment. Second, the literature on judgment and decision aids noted that individuals' confidence may increase when presented with more information. However, the increase in confidence does not necessarily imply an increase in accuracy. The difference between aided and clinical judgment is that the set of cues for aided judgment also include the predictions and standard deviations from regression models. Because the set of cues is larger for the aided judgmental approach, the effects on the auditors' confidence will also be examined.

The second set of hypotheses investigates the effects of the judgmental approaches on auditors' confidence. The additional cues in the aided judgment are anticipated to increase the auditors' confidence. Further, using arguments similar to those presented for the hypotheses for performance, auditors from the more structured firm approach are anticipated to increase their confidence with the combined approach. The hypotheses concerning confidence are:

H_{2a}: Auditors using the decision aids will express higher confidence than auditors using clinical judgment alone.

H_{21a}: Audit firm technology will affect the relative confidence of the auditors receiving aided and clinical analytical review judgment.

H_{22a}: Environmental factors will affect the relative confidence of the auditors receiving aided and clinical analytical review judgment.

Summary

This chapter has provided a summary of six relevant areas: analytical procedures, human information processing research in auditing, statistical models as analytical procedures, forecasting, judgment and decision aids, and audit technology. Each contributed to the formulation of the hypotheses.

The first two disciplines establish the importance of the research questions. The study of analytical procedures explores the professional statements and studies on analytical procedures in practice. This area motivates the importance of analytical procedures during the audit planning as an effective and efficient audit tool. Reviewing the human information processing research in auditing indicated that auditors apply heuristics in their judgment, including analytical review judgments. The bias which can be associated with heuristics indicates the need to investigate various judgmental approaches to the analytical review judgment.

The latter four disciplines (statistical models as analytical procedures, forecasting, judgment and decision aids, and audit technology) contributed to the development of the research questions and extended the considerations for the experiment. The review of statistical models as

analytical procedures indicates that regression models are a common element of the comparisons and are potential aids for use as an analytical procedure. The forecasting literature suggests that a combined approach may outperform the models and clinical judgment alone, and indicates that the relative abilities of the models are dependent upon environmental factors. Research on judgment and decision aids demonstrated that a statistical model has the potential to improve auditors' extrapolation and analytical review judgment. This literature also suggests that the training of auditors must be considered in the study, and that the auditors' confidence also needs to be investigated. The studies on audit firm technologies indicate that there is a difference in the degree of structure in Big Eight CPA firms' audit processes which affected their reaction to proposed auditing standards and which may affect the usefulness of specific audit tools.

The potential of analytical procedures to increase the effectiveness and efficiency of auditing warrants further study. The literature which is reviewed in this chapter motivates the investigation of the possibility of combining a statistical approach (regression models) with auditors' analytical review judgment. The methodology used to test the hypotheses concerning the relative merits of this combined approach will be discussed in the next chapter.

CHAPTER III

METHODOLOGY

Introduction

This chapter is divided into four main sections: (1) the subjects for the experiment and the modifications resulting from the pre-test, (2) the construction of the simulated companies and a description of the task instrument, (3) the research design, and the independent and dependent variables, and (4) the formal model.

Subjects

Subjects for this experiment were from two Big Eight CPA firms representing structured and unstructured audit technologies (see Kinney, 1986). The managing partner in charge of each local office was contacted about the project. These partners designated a contact person at each office, who in turn selected auditors who were familiar with analytical procedures. The contact person was responsible for the distribution and return of the task materials. Forty-eight instruments, each with four simulated companies, were distributed to each firm. All forty-eight instruments were returned by the more structured firm; while forty-four instruments were returned by the less structured firm (90%).

As discussed later, one subject's responses were deleted in total plus 13 additional cases from selected subjects.

Table 1 provides a summary of the subjects' background information. Three differences were noted between the subjects from the structured firm and the subjects from the unstructured firm:(1) Over twenty percent of the subjects from the structured firm had more than five years experience, but no subject from the unstructured firm had more than five years of experience. (2) About forty percent of the subjects from the structured firm were at a supervisor rank or higher, while only seven percent of the subjects from the unstructured firm were at the supervisor rank or higher. (3) Seventy-one percent of the subjects from the structured firm were certified, while eighty-eight percent of the subjects from the unstructured firm were certified. These differences may affect the results of the testing of firm structure.

Although the range of experience was greater for the subjects from the firm representing the more structured approach, approximately half of the subjects from both firms had three or four years of experience. Approximately 90 percent of the subjects from both firms were ranked at the senior level or above.

Pretest Subjects and the Results

Four CPAs from a Big Eight CPA firm in the Midwest were used to pretest the task materials (Blocher, et al., 1983).

Table 1.

Participants' Descriptive Background
Frequency Percentages

Years of Experience	Firm 1 [n = 48] Percent	Firm 2 [n = 43] Percent
less than 3	16.7	23.3
3	25.0	18.6
4	22.9	30.2
5	14.6	27.9
6	6.3	0.0
more than 6	14.6	0.0
average	4.5 years	4.1 years
Rank	**Percent**	**Percent**
Staff	8.3	11.6
Senior	52.1	81.4
Supervisor	39.6	7.0
Certification	**Percent**	**Percent**
Yes	70.8	88.4
No	29.2	11.6

The main focus was to obtain feedback on the clarity of the directions and on the realism of the task instrument. The major concerns in the original instrument were the time required to complete the task and the understanding of standard deviations. Each task was reduced to four cases and additional information about the relationship between standard deviations and confidence intervals was included both in the cover letter and in the task materials. Finally, the materials were reviewed by a manager from the

local office of each participating firm before the tasks were distributed.

The Simulation

The task materials were developed from macroeconomic data (Business Statistics: 1984, 1985) using simulation techniques. There were two major advantages for using simulation:(1) Simulation permits the direct measurement of the dependent variables for each analytical procedure because the environment of true values are known. With cases from field experience, the measurement of the dependent variables assumes the audited values are free of errors. (2) Simulated companies enable the experimenter to manipulate the population's parameters. These considerations support the use of simulated data to investigate the analytical procedures' relative strengths.

Simulated Companies

The degree of representativeness of the simulated companies impacts on the generalizability of the results. The underlying model of each company and its requirements must be carefully considered. For each company the following components were derived from the simulation. Over a 48-month period, the recorded values of the monthly account balance, the true values, and the industrial indices were derived for each company. The first 36 months represented the previous three years of audited account balances. They were used as the estimation period. The

last 12 months were the prediction period, during which the recorded values were seeded with material errors. Each company had recorded values, b(t), true or audited values, y(t), and correlated indices, z(t), for t = 1,2, ... 48. It was assumed that the recorded values and the true values for the account balance were equal in the estimation period. The specifics are discussed in the next section.

Construction of the Simulated Companies

The construction of the companies began by abstracting a core company composed of monthly sales figures and industrial indices from macroeconomic data. The four companies were developed by adjusting the core environment with random noise and seeded error. For the prediction period, t equal to 37 through 48, material error was seeded to obtain recorded values. The specifics of each step follow.

First, macroeconomic data provided f(t), the core function, and z(t), the industrial index for t equal to 1 through 48. A regression model was estimated for the total monthly sales (in the millions) of nondurable manufacturers and producer price indices for a four year period ending 1982. The model had an estimated R-square of approximately .85, which is a strong indication that regression techniques would be applicable, and had an estimated standard deviation of $2,858 on annual sales of $988,002.

The core function, f(t), captures the average

industrial characteristics. The macroeconomic data was also used to estimate the variance in the series used in simulating the variance of the random noise, E(t), component in the individual companies. The variance of the random noise was varied between two levels. These levels are discussed later as an independent variable.

The nondurable manufacturing industry was chosen because the general trend in sales was increasing and the industry was used in previous studies (see Biggs and Wild, 1985, and Kinney and Uecker, 1982). An increasing trend was chosen because Biggs and Wild (1985) indicated that the auditors were more likely to be affected by the conservative heuristic when there were increasing trends.

Second, four companies of comparable size were created using simulation to represent each of the four within-subject treatments. The monthly sales figures of the core function were proportionately reduced to $25 million in annual sales for the prediction period. The size was chosen to represent a sufficiently large client with which analytical procedures would be used.

Each level of random noise was added to the core function for two of the companies. A random normal generator produced standardized noise which was scaled to the $25 million size firm. The variance for random noise was also truncated at positive one and negative one, with

repeat sampling for outliers. The function to indicate the true value of each company, y_i, was:

$$y_i(t) = f(t) + E_i(t), \text{ where} \qquad (3.1)$$

$f(t)$ was a deterministic function of total sales of the nondurable manufacturing industry for 48 months;

$E_i(t)$ was random noise, which was distributed truncated $N(0,\sigma_i^2)$ with repeated sampling for outliers; and

i represented the level of random noise.

σ_i^2, as mentioned above, was based on the estimated standard deviation from the macroeconomic data. The two levels, which were arbitrarily chosen, were one-third and two-thirds of the estimated standard deviation.

The third structural component was seeding material errors. Material errors were seeded in each company during the prediction period, months 37 to 48. The size of the seeded error was the second factor of the within-subject treatments. The seeded error was manipulated at a level equal to materiality and at a level twice materiality.

Material errors were seeded randomly across the months of the estimation period. For each simulated company the material error was divided into four equal errors and was seeded randomly across the months in the estimation period. This seeding approach was selected because it represented a sufficiently difficult environment (see Knechel, 1984). The number of months that were seeded was not considered an

independent variable because the main focus of the experiment was the relative strength of various analytical procedures, and the experiment was constrained by subject availability.

The reported value, $b_{ij}(t)$, was a function of $y_i(t)$ and the seeding of material error, $M_j(t)$, for t equal to 37 to 48. This seeding generated four months, which required additional audit resources, and eight months, which did not require additional audit resources in the prediction period of each simulated company. The function used to represent the experimental companies was:

$$b_{ij}(t) = y_i(t) + M_j(t) = f(t) + E_i(t) + M_j(y), \text{ where} \qquad (3.2)$$

b(t) represented the reported monthly sales;
y(t) represented the simulated actual monthly sales;
f(t) represented the core company's monthly sales;
E(t) represented the random noise;
M(t) represented the seeded error;
i represented the level of random noise;
j represented the size of the seeded error; and
t represented the months 1 through 48.

Random noise, $E_i(t)$, was distributed normal $(0, \sigma_i^2)$ and seeded error, $M_j(t)$, was distributed:

$$M_j(t) = \begin{cases} 1/4 \times (\text{seeded error}) & \text{with P(1/3)} \\ \text{or} & \\ 0 & \text{with P(2/3)} \end{cases} \qquad (3.3)$$

The analytical procedures' ability to identify significant fluctuations was assessed by seeding the error at two levels.

To review briefly, macroeconomic data on the nondurable manufacturing industry was scaled to provide a

representative core company with annual sales of $25 million in the prediction period, the fourth year. The companies for the instrument were created by simulation to match the within-subject treatments, the size of the variance of random noise, and the size of the seeded error. Finally, the correlated indices were the actual producer price indices from the macroeconomic data.

The Task

This section discusses the setting of the task, the construction of the task booklets and the judgmental tasks.

The sales-receivable cycle was chosen because it is important to the audit and is prevalent in past research (e.g., Kinney and Uecker, 1982, and Biggs and Wild, 1985). In general, (net) sales is a major contribution on the income statement. In extending the judgmental research and the research concerning mechanical decision making procedures for analytical procedures, the sales figure represents a point of intersection. The valuation of sales figures is a part of many judgmental tasks (see Biggs and Wild, 1985 and Kinney and Uecker, 1982) and studies on analytical procedures (see Arrington, et al., 1983).

Construction of Task Booklets

For each company in the task booklet, the following information was provided in an identical manner. First, $z(t)$, the producer price indices, were provided for the previous 36 months and for the current year. The current

year's indices were also reproduced with each of the companies. Second, y(t), the past audited monthly sales figures, were provided for the estimation period, the first 36 months. The annual sales figures were also provided. These components were presented in a tabular form similar to those used by Lawrence, et al. (1985). Third, after discussion with the four pretest subjects, the level of materiality was defined as:

$$.04 \times (\text{net income}) = .04 \times (.10 \times \text{sales}) = .004 \times (\text{sales}) \quad (3.4)$$

Fourth, to be consistent with Biggs and Wild (1985), subjects were instructed to provide an upper and lower bound beyond which they would feel 95 percent confident that an additional investigation should be conducted. Fifth, z(t) was identified as the producer price index for all subjects and presented separately because it was common to all cases. In addition, the subjects who were assigned to the aided analytical review judgment treatment, received an explanation of the regression model, the model's predictions, the standard deviation of the model's predictions, and an explanation of the relationship between standard deviations and confidence intervals. (See Appendix.) Sixth, the mean monthly sales for the most recent year was provided as a common anchor for their predictions. These components briefly described the information which was provided to the subjects.

Lastly, the order of the companies had to be varied within the task booklets to prevent order effects. There were 24 combinations of the four within-subjects treatments.

Judgmental Task

Subjects received the task booklet and debriefing questionnaire from their contact person. The subjects were instructed that their responses would be anonymous and that they should work independently. They were instructed to perform each of the four cases independently, and to record the total time spent on the cases. After reading the introductory information about the task, the correlated indices, and the nature of regression models (if provided), the subjects were requested to begin the first case.

The subjects were instructed to provide the twelve expected monthly sales amounts and the bounds of the non-investigation region in the space provided. After completing the information for the four cases, the subjects were instructed to open the debriefing questionnaires.

The debriefing questionnaires contained the unaudited values of each case. As a second task, the subjects were asked whether or not an investigation was implied by the bounds of their non-investigation region. Subjects were then instructed to rate their confidence in the implied decision to investigate or not to investigate. If the reported value was outside the bounds, the subjects expressed their confidence in the need for additional

investigation. If the reported value was within the bounds, subjects expressed their confidence that additional investigation was not necessary.

In addition, at the conclusion of the task, each subject was asked to provide demographic information and express their opinion about the task in a debriefing questionnaire. Specifically, data were collected concerning firm affiliation, experience, the amount of time to complete the task booklet, and background in statistics. Other questions addressed the realism of the task, the time required to complete the task, additional data which they would have liked to have been given, and general reactions.

Research Design and Independent Variables

This section presents the research design of the experiment including the independent variables.

Factorial Design

The hypotheses were tested using a 2 x 2 x 2 x 2 fixed factorial design, with two between-subjects factors (audit structure and aided versus unaided judgment) and two within-subjects factors (the size of the variance of the random noise in the simulated companies and the size of the seeded error). (See Figure 2.) Pany and Reckers (1987) demonstrated that factorial designs using within-subjects factors could produce significant results while factorial designs with between-subjects would not produce significant results. Pany and Reckers, however, acknowledged the need

RESEARCH DESIGN

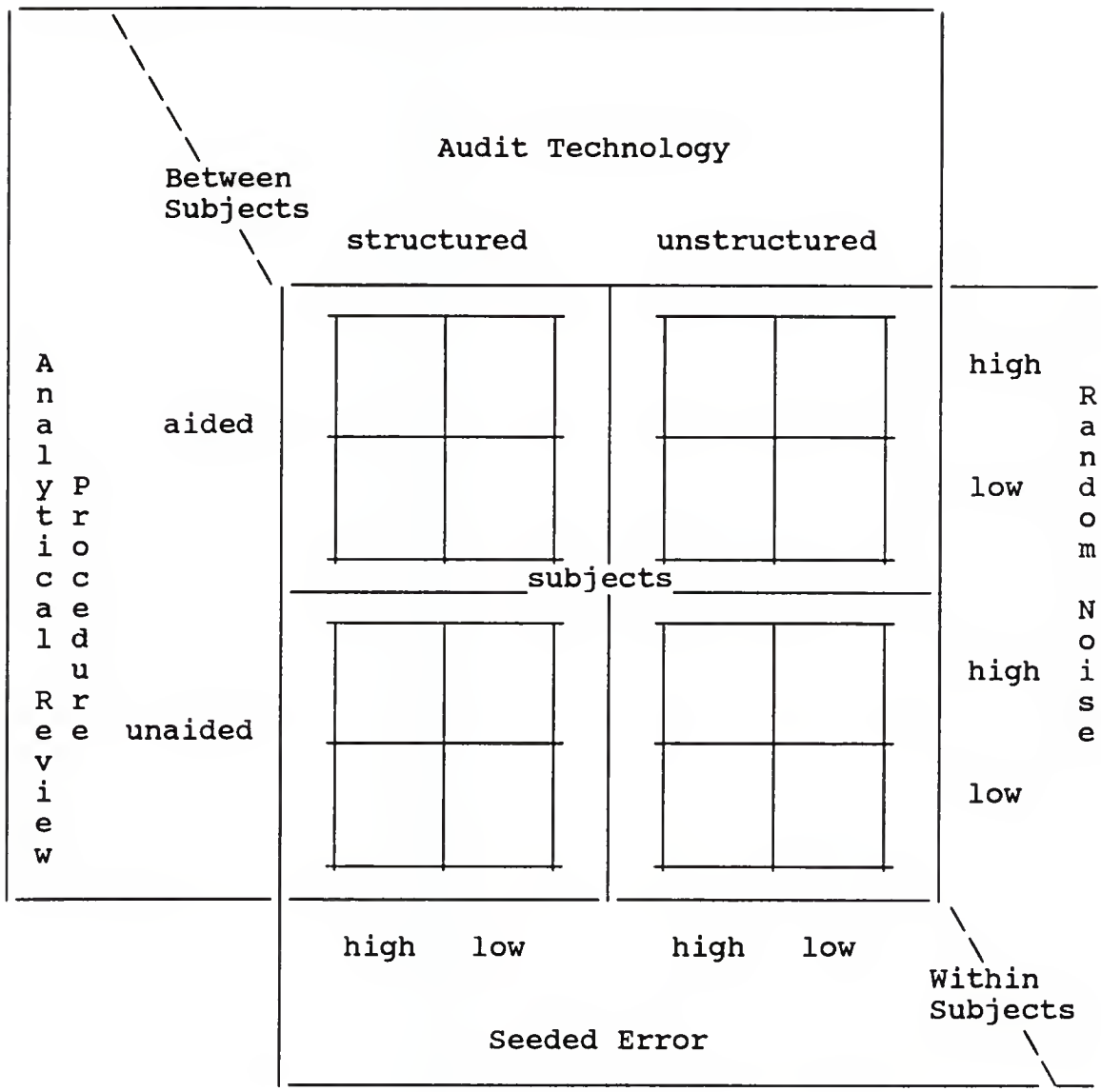


Figure 2.

for within-subjects factors when subject availability is a constraint. The within-subject factors which were used to reduce the required number of subjects, were the environmental factors. Each subject considered all four companies, but in a different order. Care was taken in the distribution of the different task booklets to prevent confounding factors with individual office effects.

Analytical Procedures

The two analytical procedures created the first between-subjects treatment. The subjects were assigned randomly to either an unaided or an aided analytical review judgment treatment. The aided judgmental approach included the predictions and the standard deviations of a regression model developed from each company's data in the prediction period.

The distinction between the treatments can be explained by Sawyer's (1966) set of judgmental approaches. The unaided approach required the subject to assimilate the past account data, to assess the time series nature of the sales figures, and to assess the correlation of these figures and their indices before predicting and evaluating the current year's data. Aided judgment allowed the regression model to organize the correlation of the data while the subjects assessed the usefulness of the model and the remaining information in the time series nature of the sales figures.

Following Lev's (1980) investigation, regression models based on the indices and the first difference of the indices were developed. The regression models based on the indices were used because they had an R-square in excess of .80 while the models based on the first difference had an R-square below .10. Furthermore, these models were evaluated using visual tests of the residuals. The models, based on the indices, more accurately approximated the assumptions of the regression models.

Audit Structure

The second between-subjects treatment was the subjects' firm audit structure. Based on the work of Cushing and Loebbecke (1986) and Kinney (1986), subjects from two firms, each representing a structured or an unstructured approach, were chosen for this factor.

Dependent Variables

Five dependent variables were extracted from the subjects' responses: the mean absolute percent error, type I error rate, Alpha, type II error rate, Beta, subjects' own assessment of their confidence in the implied decision to investigate, and their confidence in the implied decision not to investigate. Each dependent variable is considered in turn.

Mean Absolute Percent Error

Mean absolute percent error (MAPE) is a measure of the subjects' ability to estimate the true value of the account.

For each element of the underlying population there was a set of true values for the prediction period. For each month in the prediction period, the subject provided a predicted account balance for each of the four "simulated companies". MAPE was calculated separately for each company in the task booklet as follows:

$$\sum_{t=37}^{48} = |(y(t) - PV(t))/y(t)|/12, \text{ where} \qquad (3.6)$$

PV(t) was the subject's predicted value; and
y(t) was the "simulated companies'" true value for t = 37, 38, ..., 48.

MAPE provided a measure of the ability to correctly assess the current balance. An increase in predictive accuracy should enhance the analytical procedure's ability to signal fluctuations.

MAPE was chosen for the following reasons. First, the different levels of monthly sales in the companies are controlled by dividing by the true value. Second, MAPE is not affected as much by the extreme values as by squared measurements (Lawrence, et al., 1985). Third, MAPE has been used in previous studies on analytical procedures (e.g., Kinney, 1978).

Judgmental Error Rates

The non-investigation bounds provided by the subjects indicated additional investigation when the reported value was outside the bounds. These implied investigations may or may not have been appropriate. If the current reported balance was within the bounds, no significant fluctuation

was implied. These errors in the implied decision were measured using Alpha and Beta.

Alpha represented the estimate of type I errors. The procedures signaled significant fluctuations and the need for additional audit resources when the account balance was correct in terms of the materiality criteria. For example, assume the true value is $100, the reported value is $103 and the limit on material adjustment is $5. There is no material misstatement of the account balance. Consequently, if the procedure indicates $99 as the expected value, then the procedure appropriately does not signal an investigation. However, if the procedure indicates $96 as the expected value, an unnecessary investigation or change in the audit plan would occur. This error would use extra resources, but would probably not cause a reporting error because the subsequent audit findings would re-evaluate the initial findings.

Beta represented the type II errors. The procedures failed to signal significant fluctuations when there actually was a material error. As in the above example, let the reported value be $93, the true value be $100 and the materiality criteria be $5. In this example, the reported value is materially misstated. If the procedure indicates $99, then the appropriate investigation or change in the audit plan would be indicated. However, if the procedure

indicates $96, then the auditor would have to depend on subsequent audit procedures to detect the material error.

For each subject there were 12 implied decisions per company, the within-subjects treatment. One-third of the months were seeded with material error and required additional resources. Alpha and Beta were calculated as follows:

Alpha = (number of implied investigations when there was no material error)/8 (3.7)

Beta = (number of months with material errors in which an investigation was not implied)/4 (3.8)

Confidence Measures

The last two dependent variables investigated the subjects' confidence in the implied choice of their bounds of the non-investigation region. Each subject was asked to rate confidence in the implied decision, which is indicated by comparing the reported value and the bounds of the non-investigation region. Alpha confidence represented the subjects' confidence in the implied decision to investigate further. Beta confidence represented the subjects' confidence in the implied choice not to investigate further.

After completing the task, the subjects observed the reported or unaudited monthly sales figures for each company. They were asked to determine if the reported figures were within the bounds of their non-investigation regions. For reported values which were outside the regions, additional investigations were implied. Subjects

were asked to rate their confidence for each separate company in the implied decisions indicating additional investigations and the implied decision indicating additional investigations were not necessary. These variables would assess increases in the subjects' confidence because the model's expectations and standard deviations were included in the information provided. Five dependent variables were assessed from the subjects' responses.

Statistical Model

The hypotheses concerning auditors' judgment and auditors' confidence were tested using the multivariate analysis of variance (MANOVA) model because the dependent variables were anticipated to be correlated. The first hypotheses investigated the effects of the independent variables on auditors' judgment, MAPE, Alpha, and Beta. The model was:

$$\bar{Y}_{ijkln} = u + A_i + B_j + (AB)_{ij} + C_k + D_l + (CD)_{kl} + E_{ijkln}, \text{ where}$$

$\bar{Y}_{ijkln}$ is the vector of means of the dependent variables (MAPE, Alpha, Beta) of the i^{th} firm for the j^{th} decision approach on the k^{th} level of the variance of random noise with the l^{th} level of the size of seeded error for the n^{th} subject.

A_i is the main effect of the firm affiliation (CO) of the i^{th} firm.

B_j is the main effect of the decision approach (DEC) of the j^{th} approach.

C_k is the main effect of the k^{th} level of random noise (VAR).

D_l is the main effect of the l^{th} level of seeded error (SEED).

$(AB)_{ij}$ and $(CD)_{kl}$ are the interaction effects.

E_{ijkln} is the experimental error.

A similar MANOVA model was used to test confidence. ($\bar{Y}_{ijkln}$ is the vector {Alpha confidence, Beta confidence}). The actual tests will be presented in the next chapter.

Summary

To summarize the methodology section, the main components of the research method are as follows. The research design was a 2 x 2 x 2 x 2 fixed, factorial design with firm audit technologies and analytical procedures as between-subjects treatments, and the size of the variance of the random noise and the size of the seeded error as the within-subjects treatments. Four "simulated companies" were used in the experimental instrument. The order of the cases companies was varied in the different task booklets. Each "simulated company" had 36 months in the estimation period and 12 months in the prediction period. The elements were described to the subjects as 48 correlated indices, 36 past audited values and 12 current reported values. The dependent variables measured the ability to predict true values, the error rates for the decision about committing additional resources in the audit plan, and the self-reported confidence of the subjects concerning the implied decisions.

CHAPTER IV

RESULTS

Chapter IV reports the findings of the experiment. This chapter contains three sections: (1) auditor performance hypotheses, (2) auditor confidence hypotheses, and (3) subject background data. The first section examines the relative merits of aided analytical review judgment, the effects of audit firm's technology and the effects of environmental factors. This section also includes a discussion of the reasons and the number of responses which were eliminated. The second section presents the statistical tests of these effects on auditors' confidence. The third section presents descriptive data and statistical tests on background information about the subjects.

Auditor Performance Hypotheses

The first task assessed the relative merits of aided and clinical analytical review judgment and the effects of the environmental factors. Subjects provided the expected monthly sales figures and the bounds of the non-investigation region, which were used to measure performance. Three measures of judgmental performance were derived from the responses from the first task: (1) mean absolute percent error (MAPE), a measure of forecasting

accuracy, (2) Alpha, a measure of the type I error rate, and (3) Beta, a measure of the type II error rate. This section contains: (1) subject response rate, (2) analysis of auditor performance, (3) results of auditor performance, and (4) descriptive data relating to these measures.

Subject Responses

Before discussing the findings of this dissertation the response rate and the inconsistent responses need to be noted. First, 92 out of 96 (95.8 percent) of the task instruments were returned. Second, consistency checks were applied to the responses to eliminate inconsistent responses. Two consistency checks were used: (1) indicating a predicted value which was outside the bound of the non-investigation region, and (2) incorrectly responding to the question about implied investigation in the second task in the post-experimental questionnaire. If there was an inconsistent month for a case, the entire twelve months of the case company were eliminated. Most inconsistencies appear to have been caused by transposition of numbers. As a result of the consistency checks only one subject was completely eliminated and 13 individual cases (5% of the total cases) were eliminated.

Analysis of Auditor Performance

Three measures of judgmental performances were derived from subject responses (MAPE, Alpha, and Beta). Hypotheses on judgmental performance were tested using multivariate

analysis of variance (MANOVA). The model tested was presented in Chapter III. Analysis of variance (ANOVA) was applied to each of the dependent variables individually to further investigate the cause of any significant effect (Tatsuoka, 1971).

One of the assumptions of the MANOVA model is that the error terms are normally distributed. When the MANOVA was first run using MAPE, Alpha, and Beta, the error terms were not normally distributed. Four different adjustments were considered: (1) standard transformations of the dependent variables (e.g. logarithmic transformations, square root transformation, and arc sin transformations), (2) rank transformation of the dependent variables (Iman and Conover, 1979), (3) Kruskal-Wallis non-parametric tests and (4) unweighted means analysis (Lindquist, 1953).

The first three alternatives were eliminated for the following reasons: First, each of the transformations failed to improve the distribution of the error terms. Second, the rank transformation would require an indirect interpretation. Third, the Kruskal-Wallis tests assume that the dependent variables were not correlated. Thus, with the cells of approximately 20 observations each, an unweighted means analysis was applied.

Results of Auditor Performance

The means of the three dependent variables, MAPE, Alpha, and Beta, were first tested for any correlations

which might exist among them. This comparison determined whether to use multivariate analysis of variance (MANOVA) for dependent variables jointly or analysis of variance (ANOVA) for each dependent variable separately. The results are presented in Table 2.

Table 2.
Pearson Correlation Between the Means of MAPE, Alpha, and Beta
(probability)

n=16

	Alpha	Beta
MAPE	-0.09458 (.7275)	-0.44567 (.0836)
Beta	-0.62925 (.0090)	

Pearson's test for correlation indicated that two correlations were significant: Alpha and Beta ($p<.0090$), and MAPE and Beta ($p<.0836$). The negative correlation between Alpha and Beta indicated that a decrease in one error rate was associated with an increase in the other error rate. The negative correlation between Beta and MAPE indicated that a decrease in the forecasting error was associated with an increase in the type II error rate. This negative correlation could be explained by an increased non-investigation interval being associated with increased accuracy.

Table 3.

MANOVA Table for Judgmental Performance Measures

Independent Variable	Error Matrix	Approximate F based on Pillai's Trace	[df]	(p)
CO*DEC	ERROR	7.66	[3,7]	(.0129)
DEC	ERROR	5.87	[3,7]	(.0252)
CO	ERROR	12.68	[3,7]	(.0032)
VAR*SEED	ERROR	0.08	[3,7]	(.9668)
VAR	ERROR	1.47	[3,7]	(.3034)
SEED	ERROR	0.69	[3,7]	(.5851)

The results of the MANOVA (Table 3) indicated that three effects were significant: (1) the interaction effect between audit firm technology and decision approach (CO*DEC) ($F=7.66$, $p<.0129$), (2) the main decision approach effect (DEC) ($F=5.87$, $p<.0252$), and (3) the main audit firm technology effect (CO) ($F=12.68$, $p<.0032$).

To further investigate the significant effects of the MANOVA results, individual ANOVA's were tested (Tatsuoka, 1971). The results are presented in Table 4.

For each dependent variable, the overall ANOVA model had a significant F-statistic and significant effects. MAPE had significant CO*DEC effect ($F=18.18$, $p<0.0019$), DEC effect ($F=9.81$, $p<0.0121$), and CO effect ($F=20.19$, $p<0.0015$). Alpha had significant DEC effect ($F=2.09$,

Table 4.
ANOVA Results for the cell means of MAPE, Alpha, and Beta

MAPE:

Treatment	Sum of Squares	df	Mean Square	F	p<F
CO	1586.2298	1	1586.2298	20.19	0.0015
DEC	770.4788	1	770.4788	9.81	0.0121
VAR	37.8533	1	37.8533	0.48	0.5051
SEED	205.9943	1	205.9943	2.62	0.1399
CO*DEC	1477.8285	1	1477.8285	18.18	0.0019
VAR*SEED	3.9701	1	3.9701	0.05	0.8272

Overall F = 8.66 R^2 = 0.852354 p<0.0025

Alpha:

Treatment	Sum of Squares	df	Mean Square	F	p<F
CO	0.0224	1	0.0224	16.41	0.0029
DEC	0.0111	1	0.0111	2.09	0.0192
VAR	0.0019	1	0.0019	1.40	0.2675
SEED	0.0003	1	0.0003	0.24	0.6355
CO*DEC	0.0002	1	0.0002	0.15	0.7073
VAR*SEED	0.0002	1	0.0002	0.15	0.7044

Overall F = 4.41 R^2 = 0.746057 p<0.0236

Beta:

Treatment	Sum of Squares	df	Mean Square	F	p<F
CO	0.0076	1	0.0076	5.26	0.0474
DEC	0.0280	1	0.0280	19.27	0.0017
VAR	0.0012	1	0.0012	0.86	0.3781
SEED	0.0007	1	0.0007	0.45	0.5181
CO*DEC	0.0247	1	0.0247	17.03	0.0026
VAR*SEED	0.0000	1	0.0000	0.01	0.9425

Overall F = 7.15 R^2 = 0.826524 p<0.0050

$p<0.192$) and CO effect ($F=16.41$, $p<0.0029$). Beta had significant CO*DEC interaction effect ($F=17.03$, $p<0.0026$), DEC effect ($F=19.27$, $p<0.0017$), and CO effect ($F=5.26$, $p<0.0474$). Similar to the MANOVA results, there were no significant VAR*SEED, SEED, or VAR effects for any dependent variable.

Additional ANOVA tests were conducted to further investigate the significant CO*DEC interaction effects for MAPE and Beta. An ANOVA model tested the effects separately for each audit firm technology. The results are presented in Table 5.

Each overall ANOVA model for the unstructured firm had significant F-statistic and significant effects. MAPE had a significant DEC effect ($F=20.06$, $p<0.0207$). Beta had significant DEC effect ($F=77.99$, $p<0.0031$) and SEED effect ($F=6.38$, $p<.0858$).

Only the overall ANOVA model for MAPE was significant for the structured firm ($F=5.87$, $p<0.0889$). MAPE had significant DEC effect ($F=11.04$, $p<0.0450$) and VAR effect ($F=33.57$, $p<0.0102$).

Figure 3 diagrams the cell means for each dependent variable by the decision approach and the audit firm technology. The means for the environmental factors were collapsed because no significant effect was found among the MANOVA results or any of the ANOVA results (Lindquist, 1953).

Table 5.
ANOVA Results for the cell means of MAPE, and Beta, by Audit Firm Technology

Unstructured - MAPE:

Treatment	Sum of Squares	df	Mean Square	F	p<F
DEC	2191.2200	1	2191.2200	20.06	0.0207
VAR	20.0345	1	20.0345	0.18	0.6973
SEED	327.1682	1	327.1682	3.00	0.1819
VAR*SEED	24.9925	1	24.9925	0.23	0.6651

Overall F = 5.87 R^2 = 0.886672 p<0.0889

Unstructured - Beta:

Treatment	Sum of Squares	df	Mean Square	F	p<F
DEC	0.0527	1	0.0527	77.99	0.0031
VAR	0.0026	1	0.0026	3.79	0.1469
SEED	0.0043	1	0.0043	6.38	0.0858
VAR*SEED	0.0000	1	0.0000	0.01	0.8223

Overall F = 22.05 R^2 = 0.967109 p<0.0146

Structured - MAPE:

Treatment	Sum of Squares	df	Mean Square	F	p<F
DEC	57.0846	1	57.0846	11.04	0.0450
VAR	173.6316	1	173.6316	33.57	0.0102
SEED	4.8828	1	4.8828	0.94	0.4029
VAR*SEED	61.1065	1	61.1065	11.82	0.0413

Overall F = 5.87 R^2 = 0.886672 p<0.0889

Structured - Beta:

Treatment	Sum of Squares	df	Mean Square	F	p<F
DEC	0.0000	1	0.0000	0.03	0.8751
VAR	0.0000	1	0.0000	0.00	0.9894
SEED	0.0009	1	0.0009	0.50	0.5306
VAR*SEED	0.0000	1	0.0000	0.00	0.9581

Overall F = 0.13 R^2 = 0.150721 p<0.9596

Diagram of Cell Means by Decision Approach
and Audit Firm Technology

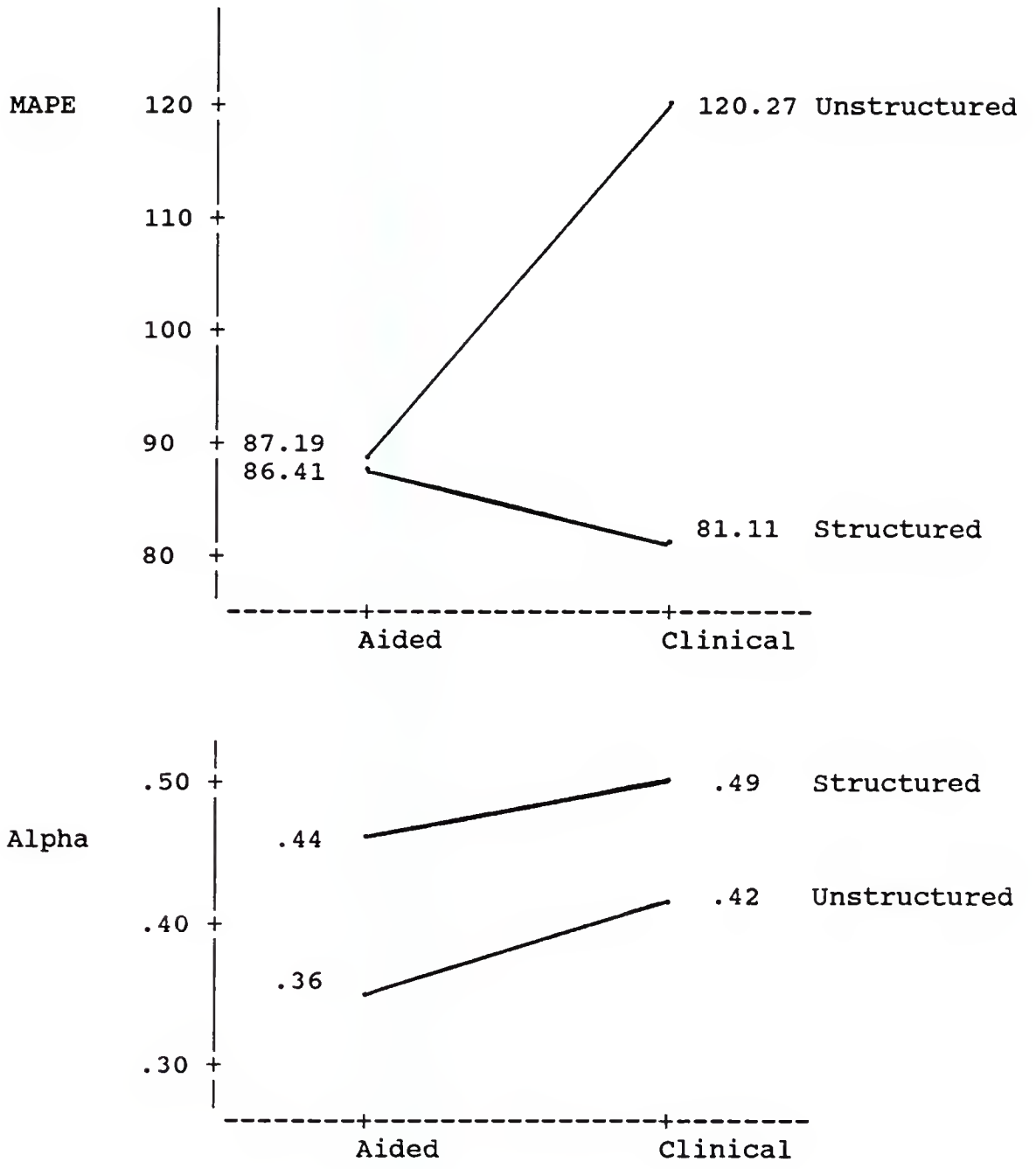


Figure 3.

Diagram of Cell Means by Decision Approach
and Audit Firm Technology

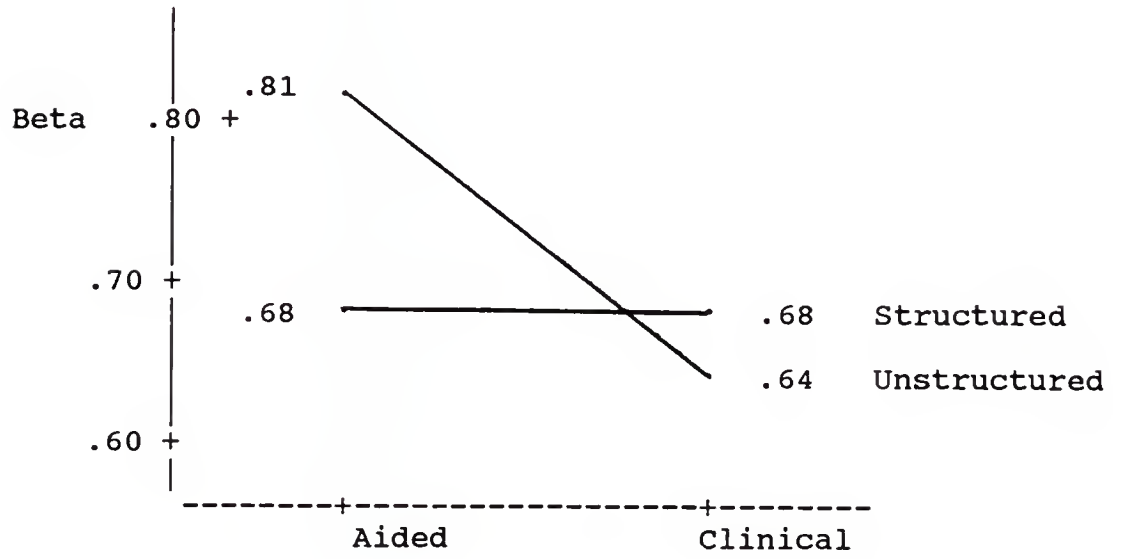


Figure 3. (continued)

The first set of hypotheses indicated that aided analytical review judgment would outperform clinical analytical judgment. The ranking was anticipated to be affected by the firm's audit technologies and by the environmental factors. The statistical evidence supports the first hypotheses that aided analytical review judgment would outperform clinical analytical review judgment for Alpha. However, the evidence for Alpha did not support the sub-hypotheses that there would be an interaction effect with the audit firms' technologies or that there would be any effects from the environmental factors. The statistical evidence indicates that there is an interaction between

decision approaches and audit firm technologies for both MAPE and Beta. For the structured firm, aided analytical review judgment was associated with an increase in the forecasting error, but a lower type I error rate, and no change in the type II error rate. For the unstructured firm, aided analytical review judgment was associated with a decrease in the forecasting error, a higher type II error rate and a lower type I error rate. The final resolution of the choice of decision approach for the unstructured firm will depend on its loss function and the interaction of analytical procedures in the unstructured firm's audit planning.

The second hypothesis suggested that there would be an interaction between the manipulated factors, the size of the variance of random noise and the size of the seeded error. There was no significant effect for the MANOVA or the ANOVAs to support this sub-hypothesis. The lack of any significant effects from the environmental factors may have been caused by one or both of the following reasons. First, the manipulations may not have been great enough to cross the subjects' thresholds. Second, the environmental factors considered in this experiment did not affect auditors' judgment.

Auditor Confidence Hypotheses

The second task measured the subjects' confidence in their implied decisions about additional investigations.

Subjects were asked to report their confidence in the implied decisions for each company in the task booklet. The results for the second task are presented in the following sections: (1) analysis of auditor confidence, (2) results of auditors confidence, and (3) descriptive data relating to these measures.

Analysis of Auditor Confidence

This section discusses the analysis of the auditors' confidence in the implied decisions about additional investigation of each company's monthly sales figures. Two measures were derived from the subjects' responses: alpha confidence (ALPHAC), which reported subjects' confidence in the implied decision for additional investigation, and beta confidence (BETAC), which reported subjects' confidence in the implied decision not to investigate further. The analysis proceeded similarly to the analysis of auditors' performance, with the exclusion of the MANOVA model.

Results of Auditor Confidence

An unweighted means analysis of ALPHAC and BETAC was performed for the same reasons discussed in reference to the results of auditors' performance. The means of the two dependent variables, ALPHAC and BETAC, were first tested for correlations between the two variables. The results are presented in Table 6.

Pearson's test for correlation did not yield a significant correlation between the mean cell responses for

Table 6.
Pearson Correlation Between the Means of ALPHAC, and BETAC
(probability)

n=16

	ALPHAC
BETAC	-0.09293 (.7321)

ALPHAC and BETAC. This suggested that on average the subjects held independent levels of confidence for the two different decisions. The treatment effects (CO*DEC, CO, DEC, VAR*SEED, VAR, and SEED) were tested using ANOVA models for each dependent variable. Table 7 presents the results.

The overall ANOVA model was significant for BETAC ($F = 3.18$, $p<0.0360$). The ANOVA model for BETAC indicated two significant effects: CO*DEC ($F=4.55$, $p<0.0617$) and DEC ($F=2.78$, $p<0.0076$). The ANOVA model for ALPHAC indicated that the CO effect was significant ($F = 5.33$, $p<0.0463$), but the overall model was not significant. As with auditors performance, there were no significant effects associated with VAR*SEED, VAR or SEED.

Additional ANOVA models for each audit firm technology were tested to further explore the significant CO*DEC interaction effect for BETAC. Only the overall ANOVA model for the unstructured firm had a significant F-statistic and significant effects. BETAC had significant DEC effect

Table 7.
ANOVA Results for the cell means of ALPHAC and BETAC

ALPHAC:

Treatment	Sum of Squares	df	Mean Square	F	p<F
CO	0.1870	1	0.1870	5.33	0.0463
DEC	0.0390	1	0.0390	1.11	0.3192
VAR	0.0046	1	0.0046	0.13	0.7269
SEED	0.0086	1	0.0086	0.24	0.6332
CO*DEC	0.0352	1	0.0352	1.00	0.3430
VAR*SEED	0.0068	1	0.0068	0.19	0.6700

Overall F = 1.34 R^2 = 0.471001 p<0.3340

BETAC:

Treatment	Sum of Squares	df	Mean Square	F	p<F
CO	0.0203	1	0.0203	2.78	0.1301
DEC	0.0856	1	0.0856	11.69	0.0076
VAR	0.0033	1	0.0033	0.45	0.5183
SEED	0.0233	1	0.0233	3.18	0.1083
CO*DEC	0.0333	1	0.0333	4.55	0.0617
VAR*SEED	0.0014	1	0.0014	0.19	0.6714

Overall F = 3.81 R^2 = 0.717348 p<0.0360

(F=6.15, p<0.0892), VAR effect (F=13.02, P<0.0366) and SEED effect (F=13.02, p<0.0366). The results are presented in Table 8.

Figure 4 diagrams the cell means for BETAC by decision approach and audit firm technology. ALPHAC was omitted because the overall ANOVA model for ALPHAC failed to produce a significant F-statistic. For the same reason mentioned in the discussion of auditors performance, the environmental factors were collapsed.

Table 8.
ANOVA Results for the cell means of BETAC,
by Audit Firm Technology

Unstructured:

Treatment	Sum of Squares	df	Mean Square	F	p<F
DEC	0.0061	1	0.0061	6.15	0.0892
VAR	0.0128	1	0.0128	13.02	0.0366
SEED	0.0128	1	0.0128	13.02	0.0366
VAR*SEED	0.0032	1	0.0032	3.25	0.1690

Overall F = 8.86 R^2 = 0.921958 p<0.0520

Structured:

Treatment	Sum of Squares	df	Mean Square	F	p<F
DEC	0.1128	1	0.1128	6.70	0.0812
VAR	0.0010	1	0.0010	0.06	0.8221
SEED	0.0105	1	0.0105	0.62	0.4872
VAR*SEED	0.0000	1	0.0000	0.00	0.9800

Overall F = 1.85 R^2 = 0.711029 p<0.3210

The results of the tests of the ANOVA models for the auditor confidence judgment are discussed in this section. The set of hypotheses suggested that subjects using the aided approach to analytical review judgment would express higher levels of confidence, and that this would be affected by the auditors' firms' technologies. The second set of hypotheses indicated that the environmental factors would also affect the level of confidence expressed by the auditors.

Diagram of Cell Means of BETAC by Decision Approach and Audit Firm Technology

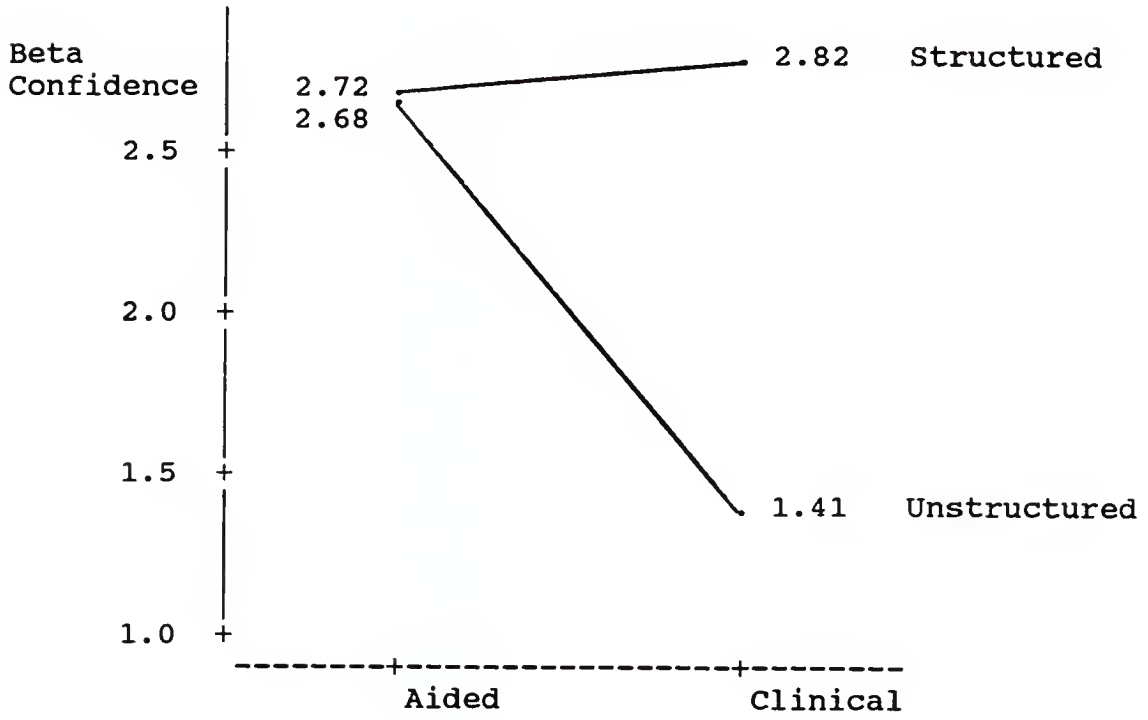


Figure 4.

The statistical evidence only supported the first set of hypotheses to some extent. The ANOVA model for the auditors' confidence in the implied decision to investigation further failed to be significant. The auditors' confidence in the implied decision not to investigate further was higher for the subjects from the unstructured firm while the auditors' confidence remained relatively unchanged for the auditors from the structured firm.

The auditors' confidence was investigated to compare the effects of regression models' information on judgment and confidence. Subjects from the structured firm decreased

their type I error rate while holding their type II error rate constant, but did not increase their confidence. Subjects from the unstructured firm, increased their forecasting accuracy, decreased their type I error rates, increased their type II error rate, and increased their confidence.

The second hypothesis suggested that the manipulated environmental factors, the size of the variance of random noise and the size of seeded error, would affect the auditors' confidence. There were no significant effects to support this hypothesis.

Subject Background Data

The subjects provided background information in the post-experimental questionnaire. This section provides the descriptive data and the results of statistical tests for: (1) the subjects' experience, (2) the subjects' perceptions of audit firm technology, (3) the subjects perceptions of analytical procedures, and (4) the subjects' perceptions of the task.

Subjects' Experience

Table 9 presents the frequency of the subjects' responses to questions about their audit experience with manufacturers of non-durable products, their educational background in regression techniques, and their experience with both mechanical and judgmental analytical procedures.

Table 9.

Frequency Distribution of Subjects' Experience by Audit Firm Technology

Have you ever audited a client, which was a manufacturer of nondurable products?

		Unstructured	Structured
a.	Yes.	67.4 %	66.7 %
b.	No.	32.6 %	33.3 %

How many courses, in which regression techniques were discussed, did you have?

		Unstructured	Structured
a.	Zero.	14.0 %	12.5 %
b.	One.	46.5 %	47.9 %
c.	Two.	32.6 %	33.3 %
d.	Three.	7.0 %	4.2 %
e.	More than three.	0.0 %	2.1 %

How many training programs have you attended which focused on the application of regression techniques in auditing?

		Unstructured	Structured
a.	Zero.	65.1 %	79.2 %
b.	One.	25.6 %	14.6 %
c.	Two.	4.7 %	6.3 %
d.	Three.	4.7 %	0.0 %
e.	More than three.	0.0 %	0.0 %

How often has a judgmental approach to analytical review been used during the audit of your clients?

		Unstructured	Structured
a.	Never.	4.7 %	0.0 %
b.	Rarely.	7.0 %	2.1 %
c.	Occasionally.	20.9 %	8.3 %
d.	Frequently.	44.2 %	72.9 %
e.	Always.	23.3 %	16.7 %

How often has a mathematical or statistical approach to analytical review been used during the audit of your clients?

		Unstructured	Structured
a.	Never.	18.6 %	25.0 %
b.	Rarely.	25.6 %	27.1 %
c.	Occasionally.	34.9 %	29.2 %
d.	Frequently.	20.9 %	18.8 %
e.	Always.	0.0 %	0.0 %

Two observations were made from these frequencies. First, the subjects from firms representing both audit technologies had a similar background in audit experience and in training with regression techniques. About two-thirds of each group had previously audited a manufacturer of non-durable goods, had one or two classes where regression techniques were discussed, and had not had any training courses in regression techniques from their firms. Second, the subjects reported varying frequencies for the application of a judgmental approach to analytical procedures. Subjects from the structured firm reported a higher frequency of a judgmental approach and a slightly lower frequency of a mathematical or statistical approach than the subjects from the unstructured firm.

These observations were investigated using Chi-Square tests. Somer's D statistic conditioned on the audit firm technology was used to augment the Chi-Square statistic for each question because at least 25 percent of the cells were less than five for most of the questions (Goodman and KruskaL, 1967). The results are presented in Table 10.

The only significant difference was the observations that subjects from the structured firm reported a higher frequency of judgmental approaches to analytical review (Chi-S`quare = 9.640, $p<0.047$). However, caution was indicated because 25 percent of the cells lacked 5 responses. The Somer's D statistic was not significant.

These results suggested that auditors from both firms had similar experience.

Subjects' Perception of Audit Firm Technology

Subjects were asked, in a series of four questions, to provide their perception of the use of formalized rules, mathematical models, and statistical approaches in their firm audit practices. Table 11 provides the frequencies of their responses to those questions.

Table 10.

Statistics on the Frequencies of Subjects' Experience

Question:	Chi-Square (prob.)	Somer's D (A.S.E.)
Have you ever audited a client, which was a manufacturer of nondurable products?	0.006 (0.937)	0.008 (0.099)
How many courses, in which regression techniques were discussed, did you have?	1.272 (0.866)	-0.008 (0.112)
How many training programs have you attended which focused on the application of regression techniques in auditing?	4.342 (0.227)	0.141 (0.094)
How often has a judgmental approach to analytical review been used during the audit of your clients?	9.640 (0.047)	-0.132 (0.111)
How often has a mathematical or statistical approach to analytical review been used during the audit of your clients?	0.729 (0.866)	0.088 (0.116)

Table 11.

Frequency Distribution Subjects' Perception of Firm Audit Technology by Firm

How often is a mathematical or statistical approach to analytical review used by your firm?

	Unstructured	Structured
a. Never.	0.0 %	8.3 %
b. Rarely.	18.6 %	35.4 %
c. Occasionally.	55.8 %	27.1 %
d. Frequently.	25.6 %	29.2 %
e. Always.	0.0 %	0.0 %

How often are mathematical or statistical rules or work-ups applied is decision-making by your firm during an audit?

	Unstructured	Structured
a. Never.	0.0 %	2.1 %
b. Rarely.	20.9 %	20.8 %
c. Occasionally.	44.2 %	39.6 %
d. Frequently.	34.9 %	37.5 %
e. Always.	0.0 %	0.0 %

How would you describe the use of statistical sampling in your firm?

	Unstructured	Structured
a. Never.	0.0 %	0.0 %
b. Rarely.	7.0 %	10.4 %
c. Occasionally.	39.5 %	33.3 %
d. Frequently.	53.5 %	50.0 %
e. Always.	0.0 %	6.3 %

How would you describe the use of structured internal control evaluations in your firm?

	Unstructured	Structured
a. Never.	0.0 %	0.0 %
b. Rarely.	2.3 %	8.3 %
c. Occasionally.	27.9 %	18.8 %
d. Frequently.	58.1 %	58.3 %
e. Always.	11.6 %	14.6 %

The distribution of the subjects' responses by audit firm indicated that the subjects' from the unstructured firm noted that the use of mechanical decision models were either

occasional or frequent for all questions. With the exception of the application of structured internal control evaluations, subjects from the unstructured firm did not respond with "never" or "always". Subjects from the structured firm had a more concentrated range of responses to the uses of mechanical decision approaches, in reference to internal control evaluation and sampling, than to their general use or their use as analytical procedures.

These observations were investigated in the same manner as the responses on the subjects' experience. The statistics are summarized in Table 12.

Table 12.

Statistics on the Frequencies of Subjects' Perception of Audit Firm Technology

Question:	Chi-Square (prob.)	Somer's D (A.S.E.)
How often is a mathematical or statistical approach to analytical review used by your firm?	10.628 (0.014)	0.173 (0.114)
How often are mathematical or statistical rules or work-ups applied is decision-making by your firm during an audit?	1.054 (0.788)	-0.003 (0.113)
How would you describe the use of statistical sampling in your firm?	3.287 (0.349)	-0.043 (0.109)
How would you describe the use of structured internal control evaluations in your firm?	2.464 (0.482)	-0.030 (0.107)

The Chi-Square statistics only indicated one significant difference. The responses about the use of mathematical of statistical approaches in analytical procedures was significant (Chi-Square = 10.628, $p<0.014$). Somer's D was also recommended for this question. The Somer's D was not significant.

Subjects' Perception of Analytical Procedures

Two questions concerned the subject opinions of analytical procedures and the task. Table 13 provides the frequencies of subject responses to questions about the effectiveness and efficiency of analytical procedures, and Table 14 summarizes the statistics on their responses.

Table 13.

Frequency Distribution of Subjects' Opinions on Analytical Review by Firm

How confident are you that analytical review procedures would identify a month with sales figures in error?

	Unstructured	Structured
a. To a very little extent.	9.3 %	4.2 %
b. To a little extent.	14.0 %	14.6 %
c. To some extent.	48.8 %	60.4 %
d. To a great extent.	25.6 %	20.8 %
e. To a very great extent.	2.3 %	0.0 %

How would you describe the efficiency of analytical review in auditing planning?

	Unstructured	Structured
a. No efficiency.	0.0 %	0.0 %
b. Less than adequate.	20.9 %	10.4 %
c. Barely adequate.	7.0 %	12.5 %
d. Adequate.	69.8 %	68.8 %
e. More than adequate.	2.3 %	8.3 %

The frequencies of responses to questions about the effectiveness of analytical procedures would suggest the following. First, subjects from the structured firm had a more concentrated range of responses to the question about their confidence in analytical procedure's effectiveness in identifying errors than subjects from the unstructured firm.

Table 14.

Statistics on the Frequencies of Subjects' Opinion on Analytical Review

Question:	Chi-Square (prob.)	Somer's D (A.S.E.)
How confident are you that analytical review procedures would identify a month with sales figures in error?	2.805 (0.591)	0.019 (0.111)
How would you describe the efficiency of analytical review in auditing planning?	3.823 (0.281)	-0.111 (0.098)

Second, about 70 percent of the subjects from both firms indicated that the efficiency of analytical procedures was adequate. The statistical tests did not indicate any significant difference between the subjects from the different firms in their perceptions of analytical procedure.

Subjects' Perception of the Task

Three questions were used to assess the subjects' interest in the task. The frequency distributions are

presented in Table 15, and the statistics are presented in Table 16.

Table 15.

Frequency Distribution of Subjects' Opinions on the Experimental Task by Firm

How well did the cases simulate actual sales data and indices?

	Unstructured	Structured
a. To a very little extent.	9.3 %	12.5 %
b. To a little extent.	20.9 %	14.6 %
c. To some extent.	60.5 %	60.4 %
d. To a great extent.	9.3 %	12.5 %
e. To a very great extent.	0.0 %	0.0 %

Total time to complete all four cases.

	Unstructured	Structured
a. Less than one hour.	11.6 %	25.0 %
b. One hour.	30.2 %	43.7 %
c. Two hours.	21.0 %	20.8 %
d. Three hours.	13.9 %	6.3 %
e. Four hours.	16.3 %	2.1 %
f. Five hours.	4.7 %	2.1 %
g. More than five hours.	2.3 %	0.0 %

If you would like a copy of the results of this study, please include your name and mailing address.

	Unstructured	Structured
a. Yes.	27.9 %	47.9 %
b. No.	72.1 %	52.1 %

The frequency distributions indicate the following observations. First, although the cases were developed strictly with macroeconomic data, more than 70 percent of the subjects indicated the data simulated actual cases to some extent or more. Second, about 65 percent of the subjects from the structured firm spent one or two hours and 10 percent spent three or more hours on the first task,

Table 16.

Statistics on the Frequencies of Subjects' Opinion on the Experimental Task

Question:	Chi-Square (prob.)	Somer's D (A.S.E.)
How well did the cases simulate actual sales data and indices?	0.942 (0.815)	-0.038 (0.106)
Total time to complete all four cases.	11.410 (0.076)	0.357 (0.108)
If you would like a copy of the results of this study, please include your name and mailing address.	3.837 (0.050)	-0.200 (0.099)

while about half of the subject from the unstructured firm spent one or two hours and 37 percent spent three or more hours. Third, almost half of the subjects from the structured firm asked for a copy of the results where as only about one-fourth of the subjects of the unstructured firm asked for a copy. The subjects from the structured firm were faster at completing the task (Chi-Square = 11.410, $p<0.076$) and more interested in the results of the study (Chi-Square = 3.837, $p<0.050$).

Summary

Three observations can be summarized from the background information. Consistent with Blocher, et al. (1983) these subjects perceived analytical procedures as an important audit tool. Contrary to expectations, subjects

from the unstructured firm expressed more exposure to mechanical or statistical analytical procedures than subjects from subjects from the structured firm. Lastly, subjects from the structured firm were more interested in the task and completed the task more quickly.

The statistical analysis indicated that two of the observed differences were significant. The subjects from The results of this chapter demonstrate that the relative merits of aided analytical review judgment on auditors' performance and confidence is conditioned upon the structure of the audit firm's technology. Although there were significant differences among the decision approaches and the audit firms' technologies for various dependent variables, the interpretations must be conditioned upon their interaction. A review of the results are provided in the next chapter.

CHAPTER V

CONCLUSIONS AND LIMITATIONS

The purpose of Chapter V is to present the conclusions and limitation of this study. This chapter is divided into four sections: (1) auditors' performance, (2) auditors' confidence, (3) limitations, and (4) future research areas.

This dissertation investigated the relative merits of statistically aided analytical review judgment on auditors' performance and confidence. The principal focus of this study was to assess the ability of analytical review judgment, aided by regression models, to increase the effectiveness and efficiency of audit planning. The findings of this dissertation indicated that aided analytical review judgment was associated with increased forecasting accuracy for the subjects from the unstructured firm, with decreased type I error rates for all the subjects, and with increased type II error rates for subjects from the unstructured firm. Any study of judgmental performance must also investigate confidence (Goldberg, 1968 and Oskamp, 1965). This study also assessed the effects of aided analytical review judgment on auditors' confidence. The findings indicated that the subjects from

the unstructured firm were more confident in the decision not to investigate further when using the aided approach.

Auditors' Performance

The study of the effects of performance was motivated by previous research on audit judgment (see Libby, 1981), on analytical procedures (e.g. Lev, 1980) and on forecasting techniques (e.g. Makridakis, 1986). The study of the effects of audit firm technology was motivated by Cushing and Loebbecke (1986). The inclusion of environmental factors, the size of the variance of random error and the size of the seeded error, was motivated by previous research on analytical procedures (e.g. Knechel, 1984).

The first task required the subjects to provide predicted values and the bounds of a non-investigation region for an analytical review of monthly sales figures for four companies. Each subject was assigned to a group representing aided or clinical analytical review judgment, and was selected from a Big Eight firm representing a structured or unstructured audit approach. Each of the four companies was created by simulation to represent a 2 x 2 factorial for the environmental factors, the size of the variance of random noise and the size of the seeded error.

The experiment indicated four findings. First, for the structured firm, a lower type I error rate was associated with the subjects using analytical review judgment. Second, for the structured firm, there was no

significant difference between the type II error rates or the forecasting accuracy associated with the subjects using aided and unaided analytical review judgment. Third, for the unstructured firm, a lower type I error rate and a higher forecasting accuracy were associated with the subjects using aided analytical review judgment. Fourth, for the unstructured firm, a higher type II error rate was associated with subjects using aided analytical review judgment.

The tests on these effects indicated that interaction between the audit firm's technology and the decision approach was significant, while the interaction between environmental factors, the size of the variance of random noise and the size of the seeded error, was not significant. There were also significant main effects. The audit firm's technology and the decision approach effects had to be conditioned upon each other because the interaction was significant. The main effects for the environmental factors were not significant.

This dissertation: (1) provides additional evidence supporting the existence of differences in audit technologies, (2) demonstrates a methodology for preliminary studies of aided audit judgment, (3) extends the previous research on analytical procedures, and (4) extends research in forecasting by combining judgment with a statistical model.

The interaction of audit firm technology and decision approach was consistent with the work of Cushing and Loebbecke (1986) and Sawyer (1966). The findings indicate that the effects of aided analytical review judgment varied between the subjects from the unstructured and structured firm. The work of Sawyer and Cushing and Loebbecke stressed the importance of considering an individual's experience and affiliation when evaluating decision approaches. The significance of this interaction effect reinforces the need to explicitly incorporate audit firm technologies when investigating audit approaches.

The use of simulation based on macroeconomic data offers a middle ground between case studies and simulation. Although macroeconomic data does not represent any single company, it does create realistic data for audit tasks. It also allows the experimenter the opportunity to manipulate parameters. The second implication is that abstract tasks based on actual data will provide insight to the effects of the different audit approaches.

This research extended the study of analytical procedures by comparing judgmental approaches with aided approaches. Libby (1981) suggested continued research on the use of decision aids. Specifically, Lev (1980) recommended the combination of regression models and analytical review judgment. The third implication of this

research demonstrated the potential for statistically aided judgment as an analytical procedure.

Makridakis (1986) also suggested combining statistical approaches with judgment to increase forecasting accuracy. The aided analytical review judgment increased the forecasting accuracy for the subjects from the unstructured firm. The fourth implication of these results was that combining statistical models and judgment has potential for increasing forecasting accuracy.

Auditors' Confidence

The inclusion of auditors' confidence was motivated by findings that suggested that individuals' confidence could be increased without increasing their judgmental performance (see Goldberg, 1968, and Oskamp, 1965). An increase in confidence without an increase in judgmental performance would be detrimental.

The second task investigated the subjects' confidence in the implied decisions about additional investigations. When the reported value was outside the subjects' bounds of the non-investigation regions, additional investigations were implied. Subjects were asked to indicate if the reported values were within or outside of the bound, and asked to rate their confidence in the implied decision about the implied activities. Because the two tasks were distributed together and are related, the factors were assigned in the same manner.

The results of the experiment indicated the following three findings. First, although subjects using aided analytical review judgment were associated with a lower type I error rate, their responses to questions about their confidence in the implied decision for further investigations, failed to produce a significant model. Second, for the structured audit firm, there was neither a change in confidence in the implied decision not to investigate further, nor a change in the type II error rate. Third, for the unstructured firm, higher confidence was associated with subjects using the aided analytical review judgment than with subjects using clinical judgment. The increased confidence was associated with an increased error rate for the subjects from the unstructured firm.

To some extent, consistent with Kahneman and Tversky (1973), Goldberg (1968) and Oskamp (1965), the aided analytical procedure did increase confidence. However, this increase was accompanied by a decrease in some aspects of performance by the subjects from the unstructured firm. Future research on decision approaches should note the effects of the approach on both performance and confidence to avoid detrimental side effects.

Limitations

The results of the experiment are subject to the limitations of the study. Laboratory experiments by their nature do not allow direct generalizability because the

sample is not random; the task is simulated; and the task is out of context. The subjects were volunteered by managing partners of large offices of two Big Eight CPA firms. Their background information indicated very little difference among the subjects from different firms. The task was created by simulation based on macroeconomic data only. The client companies required subjects to make a judgment in a vacuum void of any other data. Further, the simulation was based on an industrial average and not any individual firm. In practice, analytical procedures are actually applied in conjunction with other audit tools. This task required a judgment in isolation from other audit procedures and evidence.

The collection of data also presented three limitations. First, responses about subjects' background and perceptions were confined to multiple choice responses. Second, the reported values had to be presented to obtain the subjects' responses about their confidence. The reported values were included in the post-experimental questionnaire to prevent the reported values from influencing the subjects' responses. Third, the number of case companies used may have caused fatigue. Given the lack of significant environmental effects, the manipulation seems unnecessary.

Future Research

The success of the aided analytical review judgment in this experiment warrants additional studies. The experimental materials were developed from simulations using average macroeconomic data. Further study, using actual data in relation to other audit procedures, is necessary to confirm the results of this study. Additional studies are also recommended to expand the scope of the study of statistical models as decision aids. Additional statistical models may be more effective for analytical review judgment, and additional audit judgments might benefit from the availability of statistical information.

The results of this experiment should contribute to the research on aided decision making in auditing by indicating: (1) the need to consider the individual firm before applying decision aids, (2) the need to investigate the effects of decision approaches in auditors' confidence and (3) the need to base case studies on actual data. Continued research, using various firms and actual cases, would incorporate more informational cues and would allow analytical procedures to operate in conjunction with other audit tools.

APPENDIX*

* The information with brackets was omitted from the Clinical judgment treatment.

THE UNIVERSITY OF KANSAS

School of Business
Lawrence, Kansas 66045-2003
(913) 864-4500

Dear Participant:

The study in which you are about to participate examines the ability of statistical models to aid analytical review judgments. The experimental tasks involve judgments concerning the reasonableness of monthly sales figures during the planning stage of the audit. Each of the four hypothetical companies in your task were developed from publicly available financial information. For each company, you will be asked to provide a predicted value as well as a range of values for each month's sales figure.

[Regression models analyze the correlation between the past sales figures and the related industrial indices. In this study a regression model was developed from the previous three years' data (i.e., audited sales figures and published indices) for each company. This model provides predicted values, based on this year's published indices, for each month's sales figure and estimates of the standard deviation (i.e., the square root of the variance) of the predicted values.]

There are no right or wrong answers to this task. In reporting the results of the study all individuals and their firms will remain anonymous. Since this study is concerned with individual judgments, it is important that you work independently of other members of your firm who are participating in this study. Finally, the format is designed to meet the needs of several firms, so it may not exactly match the procedures used by your firm.

It would be appreciated if you would complete the instrument within one week from the time you receive it. Please be sure to complete the human subject consent form and the post-experimental questionnaire at the end of the instrument. When you have completed the study, please return all materials in the attached self-addressed envelope.

Enclosed you will find:

1. Human subject consent form.
2. A task booklet containing four hypothetical companies.
3. A post-experimental questionnaire.
4. A self-addressed envelope.

If you would like a copy of the results of this study, fill in the appropriate space at the end of the post-experimental questionnaire. Thank you for your time and effort.

Sincerely,

H. Francis Bush

H. Francis Bush, CPA
Assistant Professor

Human Subject Consent Form

The School of Business supports the practice of protection for human subjects participating in research. The following information is provided so that you can decide whether you wish to participate in the present study. You should feel free to withdraw at any time.

The study is concerned with the ability of regression models to assist preliminary analytical review judgments. You will receive four hypothetical companies in which you will be asked to provide expected sales figures as well as a range of values for the months of the current year. Your anonymity will be preserved in any and all findings from this research.

Your participation is solicited, but strictly voluntary. Do not hesitate to ask any questions about the study. Be assured that your name will not be associated in any way with the research findings. I appreciate your cooperation very much.

Sincerely,

H. Francis Bush, CPA
Assistant Professor
University of Kansas

_____
Signature of subject agreeing to participate

TASK BOOKLET

INSTRUCTIONS

Your task involves four companies of comparable size in the nondurable manufacturing industry. Please treat each company independently.

You are responsible for conducting an analytical review of each company's sales figures to assess the need for additional investigation of any monthly sales figure during the planning stage.

For your information, the next page shows monthly producer price indices for the nondurable manufacturing industry for the years 19x0 through 19x3. You have no reason to expect major changes this year from recent historical relationships between sales and these indices. After the producer price indices, which apply to all four companies, specific information is provided for each company.

[The sales figures for the years 19x0 through 19x2 have been audited and are assumed to be free of material errors.

Regression models use the correlation between the monthly sales figures and the corresponding indices to predict the expected values. In this task the models analyzed the past three years' monthly data to predict the 19x3 monthly sales figures from the current years' indices. A confidence interval of plus or minus two standard deviations of the predicted value will contain the "true" value approximately 95% of the time.]

Finally, the assessed materiality based on the size of the company's sales, is provided for each case.

REQUIRED:

1. For each company, indicate your best estimate of the expected audited sales figure for each month in the current year (19x3).

2. Indicate the upper and lower bounds of the possible sales values beyond which you feel 95% confident that additional investigation should be conducted to "explain" the apparent change in sales in 19x3. In other words, if monthly sales are outside of this range, you would be 95% confident that additional investigation should be undertaken.

3. Record the total amount of time to complete all four companies.

Producer Price Indices - Nondurable Manufacturers

	January	February	March	April	May	June	July	August	September	October	November	December
19x0	213.4	216.1	219.0	222.8	225.6	227.8	232.5	235.9	241.0	244.0	246.6	249.0
19x1	253.9	260.8	265.3	267.9	270.7	271.7	275.9	279.6	279.5	282.1	284.0	284.3
19x2	293.5	296.5	301.7	304.9	305.7	306.4	306.9	306.9	306.3	305.5	304.5	304.3
19x3	306.8	307.2	305.9	304.1	304.0	306.3	308.5	308.6	307.1	306.0	306.1	305.9

Company F

Monthly Sales (000's omitted.)

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Total
19x0	$1,502	$1,613	$1,718	$1,627	$1,684	$1,751	$1,657	$1,765	$1,847	$1,850	$1,839	$1,789	$20,642
19x1	$1,811	$1,939	$2,000	$1,897	$1,894	$1,931	$1,811	$1,971	$2,066	$2,054	$2,024	$2,001	$23,399
19x2	$2,039	$2,152	$2,186	$2,138	$2,112	$2,250	$2,059	$2,172	$2,213	$2,195	$2,090	$2,026	$25,632

Average monthly sales for 19x2 were $2,136 (000's omitted).

Materiality is assessed as $8,300 per month.

19x3 Producer Price Indices

Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
306.8	307.2	305.9	304.1	304.0	306.3	308.5	308.6	307.1	306.0	306.1	305.9

[Predicted Value *(000's omitted.)

$2,153	$2,156	$2,148	$2,138	$2,137	$2,151	$2,163	$2,164	$2,155	$2,149	$2,149	$2,148

Standard Deviation **(000's omitted.)

66	66	66	66	66	66	66	66	66	66	66	66]

Expected Audited Value _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____

95% Confidence Bounds Requiring an Investigation:

Upper Bound _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____

Lower Bound _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____

[* Based on a regression model which was fitted to the indices.
** The variability of the standard deviation is caused by the variability of the indices.]

Company M

Monthly Sales (000's omitted.)

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Total
19x0	$1,486	$1,588	$1,744	$1,629	$1,701	$1,759	$1,629	$1,775	$1,831	$1,842	$1,799	$1,815	$20,598
19x1	$1,786	$1,952	$1,989	$1,871	$1,861	$1,927	$1,811	$1,939	$2,055	$2,074	$2,042	$2,015	$23,322
19x2	$2,000	$2,150	$2,188	$2,155	$2,097	$2,249	$2,053	$2,156	$2,207	$2,178	$2,117	$2,014	$25,564

Average monthly sales for 19x2 were $2,130 (000's omitted).

Materiality is assessed as $8,300 per month.

19x3 Producer Price Indices

Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
306.8	307.2	305.9	304.1	304.0	306.3	308.5	308.6	307.1	306.0	306.1	305.9

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Predicted Value *(000's omitted.)	$2,148	$2,151	$2,143	$2,133	$2,132	$2,145	$2,158	$2,159	$2,150	$2,144	$2,144	$2,143
Standard Deviation **(000's omitted.)	72	72	72	72	72	72	72	73	72	72	72	72
Expected Audited Value												
95% Confidence Bounds Requiring an Investigation: Upper Bound												
Lower Bound												

* Based on a regression model which was fitted to the indices.
** The variability of the standard deviation is caused by the variability of the indices.

Company P

Monthly Sales (000's omitted.)

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Total
19x0	$1,477	$1,647	$1,773	$1,652	$1,680	$1,751	$1,688	$1,731	$1,841	$1,852	$1,770	$1,808	$20,670
19x1	$1,804	$1,947	$1,962	$1,918	$1,869	$1,973	$1,844	$1,941	$2,013	$2,034	$2,015	$2,043	$23,363
19x2	$2,014	$2,187	$2,161	$2,117	$2,134	$2,198	$2,009	$2,211	$2,217	$2,153	$2,063	$1,975	$25,439

Average monthly sales for 19x2 were $2,120 (000's omitted).

Materiality is assessed as $8,300 per month.

19x3 Producer Price Indices

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
	306.8	307.2	305.9	304.1	304.0	306.3	308.5	308.6	307.1	306.0	306.1	305.9
[Predicted Value *(000's omitted.)	$2,137	$2,140	$2,132	$2,122	$2,122	$2,134	$2,147	$2,147	$2,139	$2,133	$2,133	$2,132
Standard Deviation **(000's omitted.)	71	71	71	71	71	71	71	71	71	71	71	71]
Expected Audited Value	______	______	______	______	______	______	______	______	______	______	______	______
95% Confidence Bounds Requiring an Investigation:												
Upper Bound	______	______	______	______	______	______	______	______	______	______	______	______
Lower Bound	______	______	______	______	______	______	______	______	______	______	______	______

[* Based on a regression model which was fitted to the indices.
** The variability of the standard deviation is caused by the variability of the indices.]

Company V

Monthly Sales (000's omitted.)

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Total
19x0	$1,535	$1,596	$1,723	$1,659	$1,667	$1,728	$1,598	$1,812	$1,879	$1,828	$1,867	$1,802	$20,694
19x1	$1,824	$1,943	$1,989	$1,836	$1,871	$1,975	$1,808	$1,969	$2,060	$2,025	$2,027	$2,062	$23,389
19x2	$1,985	$2,187	$2,199	$2,098	$2,108	$2,206	$2,031	$2,203	$2,238	$2,148	$2,070	$1,986	$25,459

Average monthly sales for 19x2 were $2,122 (000's omitted).

Materiality is assessed as $8,300 per month.

19x3 Producer Price Indices

Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
306.8	307.2	305.9	304.1	304.0	306.3	308.5	308.6	307.1	306.0	306.1	305.9

Predicted Value *(000's omitted.)											
$2,142	$2,144	$2,137	$2,127	$2,126	$2,139	$2,152	$2,152	$2,144	$2,138	$2,138	$2,137
Standard Deviation **(000's omitted.)											
78	78	77	77	77	77	78	78	78	77	77	77
Expected Audited Value											
95% Confidence Bounds Requiring an Investigation:											
Upper Bound											
Lower Bound											

* Based on a regression model which was fitted to the indices.
** The variability of the standard deviation is caused by the variability of the indices.

POST-EXPERIMENTAL QUESTIONNAIRE

I. The following information will be kept strictly confidential.
   1. CPA. Yes or No. (Circle)
   2. Position within your firm._____
   3. Number of years with your firm._____
   4. Total number of years in public accounting._____

II. Total time to complete all four cases._____

III. The following sales figures represent the current unaudited monthly sales figures for 19x3 for each of the previous cases. For each of the three cases:
   1. Compare the unaudited sales figures to the range of values which you indicated as the bounds of the non-investigation region. If the unaudited value is outside of the range, additional investigation is implied. If this is the case, circle "Y" to indicate that addition investigation is implied.

      If the unaudited value is within the range of values which you indicated as the bounds of the non-investigation region, no additional investigation is implied. If this is the case, circle "N".

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
COMPANY F:												
Unaudited Value	$1,957	$2,116	$2,126	$2,057	$2,123	$2,211	$1,991	$2,125	$2,236	$2,129	$2,067	$2,032
Implied Investigation	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N

   1. How would you rate your overall confidence in the implied decisions above for additional investigations for Company F?
      a. No confidence.
      b. Somewhat confident.
      c. Fairly confident.
      d. Quite confident.
      e. Completely confident.

   2. How would you rate your overall confidence in the implied decisions above not to have additional investigations for Company F?
      a. No confidence.
      b. Somewhat confident.
      c. Fairly confident.
      d. Quite confident.
      e. Completely confident.

COMPANY M:

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Unaudited Value	$1,944	$2,103	$2,129	$2,079	$2,117	$2,185	$1,981	$2,149	$2,200	$2,181	$2,062	$2,040
Implied Investigation	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N

1. How would you rate your overall confidence in the implied decisions above for additional investigations for Company M?
   a. No confidence.
   b. Somewhat confident.
   c. Fairly confident.
   d. Quite confident.
   e. Completely confident.

2. How would you rate your overall confidence in the implied decisions above not to have additional investigations for Company M?
   a. No confidence.
   b. Somewhat confident.
   c. Fairly confident.
   d. Quite confident.
   e. Completely confident.

COMPANY P:

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Unaudited Value	$1,946	$2,129	$2,162	$2,120	$2,146	$2,236	$1,993	$2,134	$2,212	$2,155	$2,027	$2,019
Implied Investigation	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N

1. How would you rate your overall confidence in the implied decisions above for additional investigations for Company P?
   a. No confidence.
   b. Somewhat confident.
   c. Fairly confident.
   d. Quite confident.
   e. Completely confident.

2. How would you rate your overall confidence in the implied decisions above not to have additional investigations for Company P?
   a. No confidence.
   b. Somewhat confident.
   c. Fairly confident.
   d. Quite confident.
   e. Completely confident.

COMPANY V:

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Unaudited Value	$1,921	$2,120	$2,199	$2,012	$2,053	$2,197	$1,998	$2,172	$2,191	$2,066	$2,071	$2,025
Implied Investigation	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N

1. How would you rate your overall confidence in the implied decisions above for additional investigations for Company V?
   a. No confidence.
   b. Somewhat confident.
   c. Fairly confident.
   d. Quite confident.
   e. Completely confident.

2. How would you rate your overall confidence in the implied decisions above not to have additional investigations for Company V?
   a. No confidence.
   b. Somewhat confident.
   c. Fairly confident.
   d. Quite confident.
   e. Completely confident.

IV. Circle the appropriate response.

1. How confident are you that analytical review procedures would identify a month with sales figures in error?
   a. To a very little extent.
   b. To a little extent.
   c. To some extent.
   d. To a great extent.
   e. To a very great extent.

2. How would you describe the efficiency of analytical review in audit planning?
   a. No efficiency.
   b. Less than adequate.
   c. Barely adequate.
   d. Adequate.
   e. More than adequate.

3. How often has a judgmental approach to analytical review been used during the audit of your clients?
   a. Never.
   b. Rarely.
   c. Occaslonally.
   d. Frequentiy.
   e. Always.

4. How often has a mathematical or statistical approach to analytical review been used during the audit of your clients?
   a. Never.
   b. Rarely.
   c. Occasionally.
   d. Frequently.
   e. Always.

V. Circle the appropriate response.

1. How often is a mathematical or statistical approach to analytical review used by your firm?
   a. Never.
   b. Rarely.
   c. Occasionally.
   d. Frequently.
   e. Always.

2. How often are mathematical or statistical rules or workups applied in decision-making by your firm during an audit?
   a. Never.
   b. Rarely.
   c. Occasionally.
   d. Frequently.
   e. Always.

3. How would you describe the use of statistical sampling in your firm?
   a. Never.
   b. Rarely.
   c. Occasionally.
   d. Frequentiy.
   e. Always.

4. How would you describe the use of structured internal control evaluations in your firm?
   a. Never.
   b. Rarely.
   c. Occasionally.
   d. Frequently.
   e. Always.

VI. Respond, as indicated, to the following questions.

1. How well did the cases simulate actual sales data and indices?
   a. To a very little extent.
   b. To a little extent.
   c. To some extent.
   d. To a great extent.
   e. To a very great extent.

2. Which form of information was easier to assess, the correlation between the sales figures and the indices or the trend in the sales figures over time?

3. What other information would you have liked to have incorporated into the cases?

4. Explain the difficulties, if any, you had with the tasks and make any other comments you feel appropriate about the task.

VII. Circle the appropriate response.

1. Have you ever audited a client, which was a manufacturer of nondurable products?
   a. Yes.
   b. No.

2. How many courses, in which regression techniques were discussed, did you have?
   a. Zero.
   b. One.
   c. Two.
   d. Three.
   e. More than three.

3. How many training programs have you attended which focused on the application of regression techniques in auditing?
   a. Zero.
   b. One.
   c. Two.
   d. Three.
   e. More than three.

VIII. If you would like a copy of the results of this study, please include your name and mailing address.

_____
_____
_____
_____
_____
_____
_____

BIBLIOGRAPHY

Akresh, A., and W. Wallace, "The Application of Regression Analysis for Limited Review and Audit Planning," Symposium on Auditing Research IV, (University of Illinois Trustees, 1981).

Albrecht, W. S., "Toward Better and More Efficient Audits," Journal of Accountancy, (December, 1977), pp. 48-50.

_____, and J. C. McKeown, "Toward an Extended Use of Statistical Analytical Reviews in the Audit," Symposium on Auditing Research II, (University of Illinois Trustees, 1977), pp. 53-69.

American Institute of Certified Public Accountants, Statement on Auditing Standards, No.56, Analytical Procedures, (New York: AICPA, 1988).

_____, Codification of Statements on Auditing Standards, Nos. 1-39, (New York: AICPA, 1982).

Arens, A. A. and J. K. Loebbecke, Auditing: An Integrated Approach, second edition, (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1980).

Armstrong, J. S., "Relative Accuracy of Judgmental and Extrapolative Methods in Forecasting Annual Earnings," The Journal of Forecasting, Volume 2, (1983), pp. 437-447.

Arrington, C. E., W. Hillison and R. C. Icerman, "Research in Analytical Review: The State of the Art," Journal of Accounting Literature, (1983), pp. 151-185.

Ashton, R. H., "Research in Audit Decision Making: Rationale, Evidence and Implication," an unpublished manuscript, New York University, (January, 1983).

Barrows, H. S., and K. Bennett, "The Diagnostic (Problem Solving) Skill of the Neurologist: Experimental Studies and Their Implications for Neurological Training," Archives of Neurology, Volume 26, (1972), pp. 273-277.

Biggs, S. F., T. J. Mock, and P. R. Watkins, "A Descriptive Study of Auditors' Use of Analytical Review in Audit Program Design," an unpublished manuscript, (1985).

Biggs, S. F. and J. J. Wild, "An Investigation of Auditor Judgment in Analytical Review," The Accounting Review, (October, 1985), pp. 607-633.

_____, "A Note on the Practice of Analytical Review," Auditing: A Journal of Practice and Theory, (Spring, 1984), pp. 68-79.

Blocher, E., R. S. Esposito and J. J. Willingham, "Auditors' Analytical Review Judgments for Payroll Expenses," Auditing: A Journal of Practice and Theory, (Fall, 1983), pp. 66-90.

Box, G. P., and G. M. Jenkins, Time Series Analysis: Forecasting and Control, revised edition, (San Francisco, California: Holden-Day, 1976).

Chatfield, C., The Analysis of Time Series: An Introduction, second edition, (London, England: Chapman and Hall, 1980).

Coakley, J. R., Analytical Review: A Comparison of Procedures and Techniques Used in Auditing, Ph. D. Dissertation, The University of Utah, (1982).

Collins, W. A., and W. S. Hopwood, "A Multivariate Analysis of Annual Earnings Forecasts Generated from Quarterly Forecasts of Financial Analysts and Univariate Time-Series Models," Journal of Accounting Research, (Autumn, 1980), pp. 390-406.

Cushing, B. E. and J. K. Loebbecke, Comparison of Audit Methodologies of Large Accounting Firms, American Accounting Association, (Sarasota, Florida, 1986).

Dawes, R. M., "A Case Study of Graduate Admission: Application of Three Principles of Human Decision Making," American Psychologist, Volume 26, (1971), pp. 180-188.

_____, and B. Corrigan, "Linear Models in Decision Making," Psychological Bulletin, Volume 81, (1974), pp. 95-106.

de Dombal, F. T., and J. C. Horrocks, "Computed Aided Diagnosis: Conclusions from an Overall Experience Involving 4,469 Patients," J. Anderson, and J. M. Forsythe (eds.), Medinfo '74, (New York: American Elsevier/North Holland, 1974).

Demski, J. S., Information Analysis, (Reading, Massachusetts: Addison-Wesley, 1972).

Edwards, W., "Conservatism in Human Information Processing," Formal Representation of Human Judgment, B. Kleinmuntz (ed.), (New York: Wiley, 1968).

_____, H. Lindman, and L. D. Phillips, "Emerging Technologies for Making Decisions," Chapter 2 in New Directions in Psychology, (New York: Holt, Rhinehart, and Winston, 1965).

Eggleton, I. R., "Intuitive Time-Series Extrapolation," Journal of Accounting Research, (Spring, 1982), pp. 68-102.

Einhorn, H. J., "Expert Measurement and Mechanical Combination," Organizational Behavior and Human Performance, Volume 7, (1972), pp. 86-106.

_____ and R. M. Hogarth, "Behavioral Decision Theory: Processes of Judgment and Choice," Journal of Accounting Research, (Spring, 1981), pp. 1-41.

_____ and R. M. Hogarth, "Confidence in Judgment: Persistence of the Illusion of Validity," Psychology Review, Volume 85, No. 5, (1978), pp. 395-416.

_____, D. N. Kleinmuntz and B. Kleinmuntz, "Linear Regression and Process-Tracing Models of Judgment," Psychology Review, Volume 86, No. 5, (1979), pp. 465-485.

Elstein, A. S., L. S. Shulman and S. A. Sprafka, Medical Problem Solving: An Analysis of Clinical Reasoning, (Cambridge, Massachusetts: Harvard University Press, 1978).

Felix, W. L. and W. R. Kinney, Jr., "Research in Auditor's Opinion Formulation Process: State of the Art," The Accounting Review, (April, 1982), pp. 245-271.

Foster, G., "Quarterly Accounting Data: Time Series Properties and Predictive Ability Results," The Accounting Review, (January, 1977), pp. 1-21.

Fryback, D. G., and J. R. Thornbury, "Informal Use of Decision Theory to Improve Radiological Patient Management," Radiology, Volume 129, (1978), pp. 385-388.

Gaumnitz, B. R., T. R. Nunmaker, J. S. Surdick and M. F. Thomas, "Auditor Consensus in Internal Control Evaluation and Audit Program Planning," Journal of Accounting Research, (Autumn, 1982, Part II), pp. 745-755.

Gibbons, M., "Regression and Other Statistical Implications for Research on Judgment Using Intercorrelated Data Sources," Journal of Accounting Research, (Spring, 1982), pp. 68-102.

Gill, P. W., D. J. Leaper, P. J. Guillou, J. R. Staniland, J. C. Horrocks and F. T. de Dombal, "Observer Variation in Clinical Diagnosis--A Computer-Aided Assessment of its Magnitude and Importance in 552 Patients with Abdominal Pains," Methods of Information in Medicine, Volume 12, (1973), pp. 108-113.

Goldberg, L. R., "Man Versus Model of Man: A Rationale, Plus Some Evidence for a Method of Improving on Clinical Inferences," Psychological Bulletin, Volume 73, (1970), pp. 422-432.

_____, "Simple Models or Simple Processes? Some Research on Clinical Judgments," American Psychologist, Volume 3, (1968), pp. 483-496.

Goodman, L. A. and W. H. Kruskal, "Measures of Association for Cross-Classification IV," Journal of American Statistical Association, Volume 67, (1972), pp. 415-421.

Hammond, K. R., "Computer Graphics as an Aid to Learning," Science, Volume 172, (1971), pp. 903-908.

Hilton, R., "Integrating Normative and Descriptive Theories of Information Processing," Journal of Accounting Research, (Autumn, 1980), pp. 477-505.

Hogarth, R. M., Judgement and Choice, The Psychology of Decision, (New York: Wiley, 1980).

Holder, W. W., "Analytical Review Procedures in Planning the Audit: An Application Study," Auditing: A Journal of Practice and Theory, (Spring, 1983), pp. 11-18.

Hollander, M. and D. A. Wolfe, Nonparametric Statistical Methods, (New York: Wiley, 1973).

_____, and S. Makridakis, "Forecasting and Planning: An Evaluation," Management Science 31, No. 10, (1981), pp. 1237-1246.

Hylas, R. and R. Ashton, "Audit Detection of Financial Statement Errors," The Accounting Review, (October, 1982), pp. 751-765.

Iman, R. L. and W. J. Conover, "The Use of the Rank Transform in Regression," Technometrics, (November, 1979), pp.499-509.

Joyce, E. J., "Expert Judgment in Audit Program Planning," Studies on Human Information Processing in Accounting, Supplement to Journal of Accounting Research, (1976), pp. 29-60.

_____, G. C. Biddle, "Anchoring and Adjustment in Probabilistic Inference in Auditing," Journal of Accounting Research, (Spring, 1981), pp. 120-145.

Kahneman, D. and A. Tversky, "On the Psychology of Prediction," Psychological Review, Volume 80, (1973), pp. 237-251.

Kaplan, R. D., "Developing a Financial Planning Model for Analytical Review: A Feasibility Study," Symposium on Auditing Research III, (Urbana-Champaign, Illinois, University of Illinois Trustees,1979), pp. 3-30.

Kendall, R. E., The Role of Diagnosis in Psychiatry, (Oxford, England: Blackwell Scientific Publications, 1975).

Kinney, W. R., Jr., "Audit Technology and Preferences for Auditing Standards," Journal of Accounting and Economics, (March, 1986), pp. 73-89.

_____, "Quantitative Applications in Auditing," Journal of Accounting Literature, (1983), pp. 187-204.

_____, "The Predictive Power of Limited Information in Preliminary Analytical Review: An Empirical Study," Supplement to Journal of Accounting Research, (1979), pp. 148-165.

_____, "ARIMA and Regression in Analytical Review: An Empirical Test," The Accounting Review, (January, 1978), pp. 48-60.

_____ and G. Salamon, "Regression Analysis in Auditing: A Comparison of Alternative Investigation Rules," Journal of Accounting Research, (Autumn, 1982, Part I), pp. 350-367.

_____ and W. C. Uecker, "Mitigating the Consequences of Anchoring and Adjustment in Auditor Judgments," The Accounting Review, (April, 1982), pp. 245-271.

Kleinmuntz, B., "The Processing of Clinical Information by Man and Machine," Formal Representation of Human Judgment, B. Kleinmuntz (ed.), (New York: Wiley, 1968).

Knechel, W. R., "A Simulation Study of the Relative Effectiveness of Alternative Analytical Review Procedures," an unpublished manuscript, University of Florida, (June, 1984).

Lawrence, M. J., R. H. Edmundson and M. J. O'Connor, "An Examination of Judgmental Extrapolation of Time Series," International Journal of Forecasting, Volume 1, (1985), pp. 25-35.

Leaper, D. J., J. C. Horrocks, J. R. Staniland and F. T. de Dombal, "Computer Assisted Diagnosis of Abdominal Pain Using 'Estimates' Provided by Clinicians," British Medical Journal, Volume 4, (1972), pp. 350-354.

Lev, B., "On the Use of Index Models in Analytical Reviews by Auditors," Journal of Accounting Research, (Fall, 1980), pp. 524-550.

Libby, R., Accounting and Human Information Processing: Theory and Applications, (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1981).

Lindquist, E. F., Design and Analysis of Experiments in Psychology and Education, (Boston: Houghton Mifflin Company, 1953).

Makridakis, S., "The Art and Science of Forecasting: An Assessement and Future Directions," Internation Journal of Forecasting, Volume 2, (1986), pp. 15-39.

_____ and R. L. Winkler, "Average of Forecasts: Some Empirical Results," Management Science, Volume 29, No. 9, (September, 1983), pp. 987-996.

_____, A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen and R. Winkler, "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition," Journal of Forecasting, Volume 1, (1982), pp. 111-153.

_____, and M. Hibon, "Accuracy of Forecasting: An Empirical Investigation," Journal of the Royal Statistical Society, Volume 142, part 2, (1979), pp. 97-145.

Mendenhall, W. and R. L. Schaeffer, Mathematical Statistics and Applications, (Massachusetts: Duxbury Press, 1973).

Montgomery, D. C., Design and Analysis of Experiments, (New York: Wiley, 1976).

Neter, J., "Two Case Studies on the Use of Regression for Analytical Review," Symposium on Auditing Research IV, (University of Illinois Trustees, 1980).

Olsen, C. L., "On Choosing a Test Statistic in Multivariate Analysis of Variance," Pyschological Bulletin, Volume 83, (1976), pp. 579-586.

Oskamp,S. , "Overconfidence in Case-Study Judgments," Journal of Consulting Psychology, Volume 29, (1965), pp. 261-265.

Pany, K. and M. J. Reckers, "Within- vs. Between-Subjects Experimental Designs: A Study of Demand Effects," Auditing: A Journal of Practice and Theory, (Fall, 1987), pp. 39-53.

Pitz, G. F. and N. J. Sachs, "Judgment and Decision: Theory and Application," Annual Review of Psychology, M. R. Rosenzweig and L. W. Porter (eds.), (Palo Alto, California: Annual Review, Inc., 1984), pp. 139-165.

Sawyer, J., "Measurement and Prediction: Clinical and Statistical," Psychological Bulletin, Volume 66, (1966), pp. 178-200.

Senior, J. R., The Development and Validation of a Computer-Based System for Testing and Teaching Clinical Competence, (Philadelphia, Pennsylvania: National Board of Medical Examiners and American Board of Internal Medicine, 1974).

Slovic, P. and S. Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing Judgments," Organizational Behavior and Human Performance, Volume 6, (1971), pp. 649-744.

Smedslund, J., "The Concept of Correlation in Adults," Scandinavian Journal of Psychology, Volume 4, (1963), pp. 165-174.

Tatsuoka, M. M., Multivariate Analysis Techniques for Educational and Psychological Research, (New York: John Wiley, and Sons, Inc., 1971).

Tuddenham, W. J., L. M. Houser, P. S. Tuddenham, R. E. Booth, Jr. and S. Mattews, "Preliminary Evaluation of Effectiveness of Logical Flow Charts in Teaching Roentgen Diagnosis," Radiology, Volume 93, (1969), pp. 17-24.

Tversky, A. and D. Kahneman, "Judgment Under Uncertainty: Heuristics and Biases," Science, Volume 185, (1974), pp. 1124-1131.

_____, "Availability: A Heuristic for Judging Frequency and Probability," Cognitive Psychology, Volume 5, (1973), pp. 207-232.

Wallace, W. A., "The Acceptability of Regression Analysis as Evidence in a Courtroom--Implications for the Auditor," Auditing: A Journal of Practice and Theory, (Spring, 1983), pp. 66-90.

Warner, H. R., A. F. Toronto, G. Veasey and R. Stephenson,"A Mathematical Approach to Medical Diagnosis: Application to Congenial Heart Disease," Journal of American Medical Association, Volume 177, (1961), pp. 177-183.

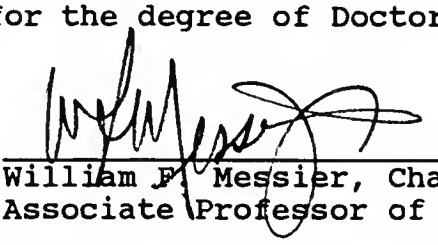
Wickens, C. D., Engineering Psychology and Human Performance, (Columbus, Ohio: Charles E. Merrill Publishing Company, 1984).

Wright, G. and K. Ng, "Three Methods in Aiding Decision Making," Clinical Psychology and Medicine: A Behavioral Perspective, Main, C. J. and W. R. Lindsay, Eds., (New York, New York, Plenum Press, 1982).

BIOGRAPHICAL SKETCH

Howard Francis Bush was born May 13, 1952, in Rutherford, New Jersey. He graduated from Monroe-Woodbury Central High School in June, 1970. Mr. Bush attended the State University of New York at Buffalo and graduated with a Bachelor of Arts in mathematics in 1973. After a year of teaching secondary mathematics in Buffalo, Mr. Bush began his graduate studies at The Ohio State University. He completed his Masters of Accountancy in June, 1977 and taught at West Virginia University and the University of Wisconsin-Milwaukee as a lecturer. Mr. Bush entered the doctoral program at the University of Florida in 1980. While at the University of Florida, Mr. Bush was the recipient of the Warner Ring Foundation Scholarship, Deloitte Haskins and Sells Doctoral Fellowship, and Outstanding Graduate Teaching Assistant Award. He married Felicia Linda DiPronio of Milwaukee, Wisconsin, in January, 1982. Mr. Bush accepted his first faculty position at the University of Kansas, Lawrence, Kansas.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully acceptable, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

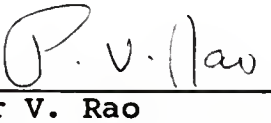


William F. Messier, Chairman
Associate Professor of Accounting

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully acceptable, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

E. Michael Bamber
Assistant Professor of Accounting

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully acceptable, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Pejaver V. Rao
Professor of Statistics

This dissertation was submitted to the Graduate Faculty of the Fisher School of Accounting in the College of Business Administration and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 1989

_____

Dean, Graduate School